Computers and Electronics in Agriculture 191 (2021) 106519

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Reliability of a commercial platform for estimating flower cluster and fruit number, yield, tree geometry and light interception in apple trees under different rootstocks and row orientations

Alessio Scalisi [a,*], Lexie McClymont [a], James Underwood [b], Peter Morton [b], Steve Scheding [b], Ian Goodwin [a,c]

[a] Tatura SmartFarm, Agriculture Victoria, Tatura, VIC 3616, Australia
[b] Green Atlas, Alexandria, NSW 2015, Australia
[c] Centre for Agricultural Innovation, University of Melbourne, Parkville, VIC 3010, Australia

ARTICLE INFO

Keywords:
Canopy size
Cartographer
Malus domestica
LiDAR
Mapping
Orchard

ABSTRACT

Modern horticulture is undergoing a rapid change with the introduction of new predictive technologies that help maximise the automation of orchard management practices. This study aimed to calibrate and validate a commercial sensorised mobile platform for the prediction of flower cluster number, fruit number and yield, tree geometry in 'ANABP-01′ apples. In addition, this work (i) modelled the relationships between tree geometry and light interception, and (ii) determined the effects of light interception, rootstock and row orientation on flower cluster number, crop load, yield and tree geometry. Results showed that predictions were very accurate after initial calibration. Flower cluster detections had an error (*RMSE*) of ~ 5 clusters / image. Fruit number and yield predictions needed independent calibration across rootstocks but errors after validation on a separate dataset were small (*RMSE* = 5 fruit / tree, and *RMSE* = 1 kg / fruit, for fruit number and yield, respectively). Orchard errors for fruit number and yield estimations were lower than 5 %. Canopy area, canopy density and canopy cross-sectional leaf area (CSLA) were all linearly related with effective area of shade (EAS, integrated daily canopy light interception) but CSLA had the most robust and stable relationship with intercepted light. Increasing CSLA led to higher flower cluster number, fruit number and yield. Row orientations and rootstocks significantly affected productive performance, tree size and geometry and light interception. The orchard heatmaps generated after data validation proved very useful to support orchard management decisions. Overall, the predictive technology demonstrated to be a valid tool to combine accurate estimates of several important fruit crop parameters (i.e. flower cluster number, fruit number, yield, tree size and geometry, and light interception) in a single platform.

1. Introduction

Modern horticulture is moving toward increased mechanisation, automation, robotics, and non-destructive sensing and monitoring. The integration of technologies that are already adopted in other industries into horticulture systems aims to increase resource use efficiency — including labour — and make orchards more profitable. For this purpose, several recent studies have focused on the application of machine learning algorithms to detect tree structures (e.g. flowers, fruit, architecture) using sensorised robots or platforms. Most of the state-of-the-art research has attempted to detect apple fruit for fruit number or yield determination, or for integration with automated harvesting machines using image segmentation, deep learning and different Convolutional Neural Networks (CNN) (Bargoti and Underwood, 2017a, 2017b; Bresilla et al., 2019; Kang and Chen, 2020; Kuznetsova et al., 2020) on images typically collected by RGB / RGB-D cameras. Underwood et al. (2016), Dias et al. (2018a, 2018b) and Wang et al. (2018) used similar machine vision approaches for almond, apple and mango flower recognition, respectively. In the case of almond, the machine image recognition was supported by LiDAR cloud points to reconstruct tree structure and assign tree geo-references when combined with GPS (Underwood et al., 2016). LiDAR sensors are a powerful tool to quickly

* Corresponding author at: 255 Ferguson Rd, Tatura, VIC 3616, Australia.
E-mail address: alessio.scalisi@agriculture.vic.gov.au (A. Scalisi).

https://doi.org/10.1016/j.compag.2021.106519
Received 29 July 2021; Received in revised form 21 October 2021; Accepted 24 October 2021
Available online 31 October 2021
0168-1699/

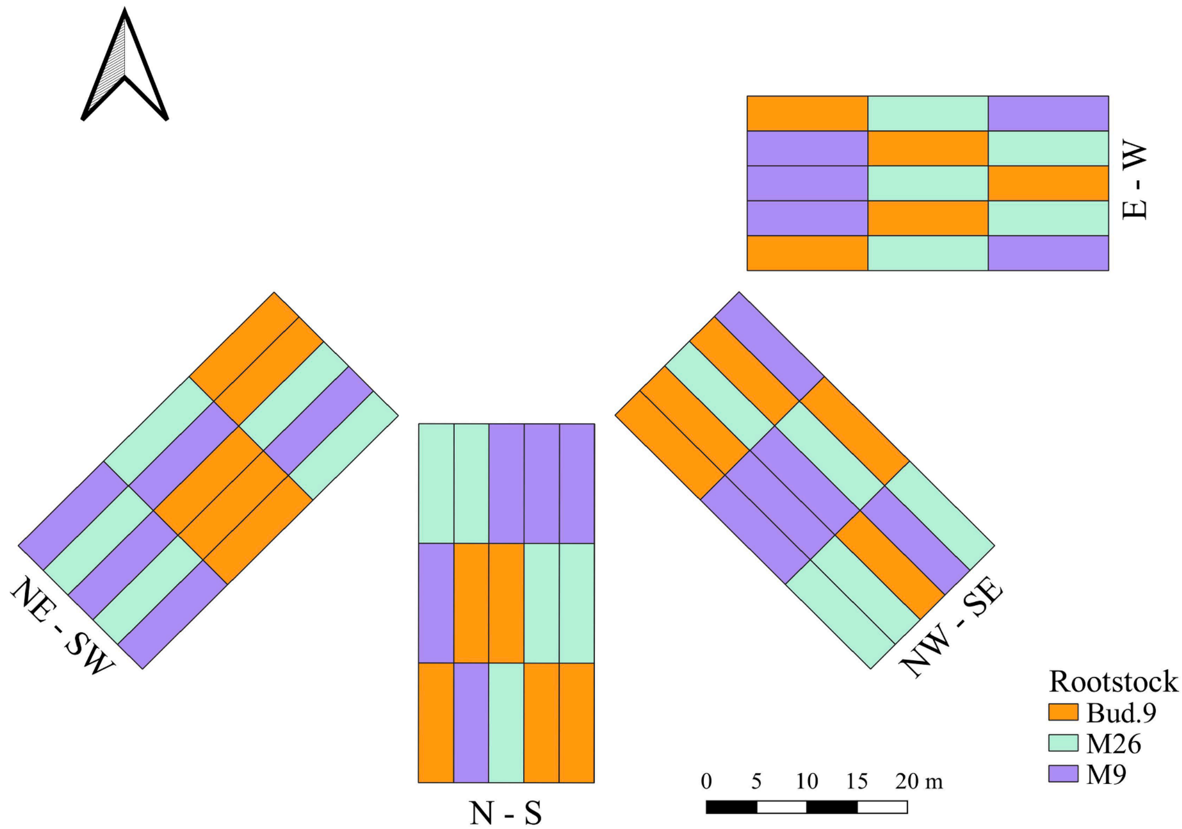

Fig. 1. Layout of the sixty experimental plots of 'ANABP-01′ apples in the Sundial orchard at the Tatura SmartFarm. Row orientations: northeast–southwest (NE – SW), north–south (N – S), northwest–southeast (NW – SE) and east–west (E – W).

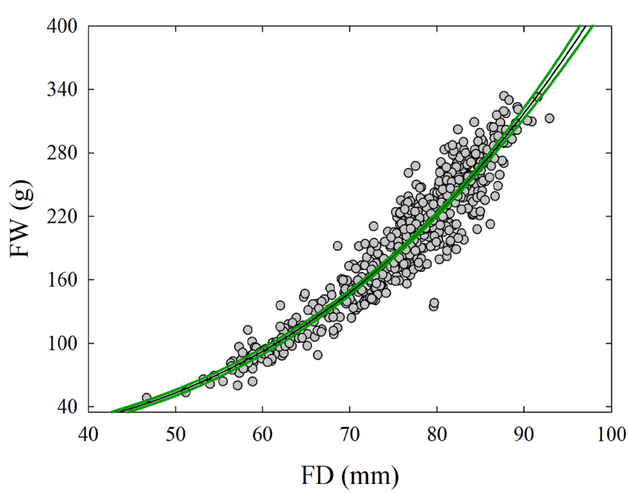

Fig. 2. Fruit weight (FW) against fruit diameter (FD) in 'ANABP-01′ apple. FW = 0.0003 * $FD^{3.04}$. Black line represents the power regression fit; green lines show the 95% confidence interval bands; standard error of the estimate = 21 g; n = 559. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

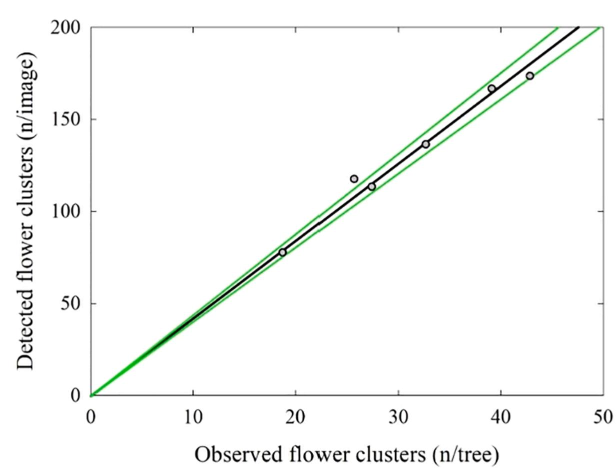

Fig. 3. Detected against observed flower clusters in 'ANABP-01′ apple trees at 50 % bloom. Black and green lines are the linear regression fit [y = 4.205 (0.069) x; *RMSE* = 5 cluster / image] and 95% confidence interval bands, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

determine canopy architecture parameters such as tree height, canopy size and canopy density and have the potential to recognise tree location, alone or combined with GPS (Underwood et al., 2015). LiDAR cloud points have also been used to model light interception, as demonstrated by Örn (2016). The same idea was applied to estimate a solar-geometric model for light interception estimation in avocado and mango trees (Westling et al., 2018; 2020), and further extended to make pruning recommendations (Westling et al., 2021).

Commercial services such as *Cartographer* (Green Atlas) use a combination of sensors (e.g. RGB cameras, LiDAR, GPS), mounted on a platform such as an electric all-terrain vehicle (ATV), to gather data while driving through orchard rows. *Cartographer* is currently available to measure the spatial distribution of fruit number in apples and to measure tree geometry parameters such as tree height, canopy area and canopy density; in addition, *Cartographer* is being tested for predictions of fruit parameters such as fruit colour, fruit size and fruit clustering. Here canopy area (m^2) represents the area of the polygon drawn around

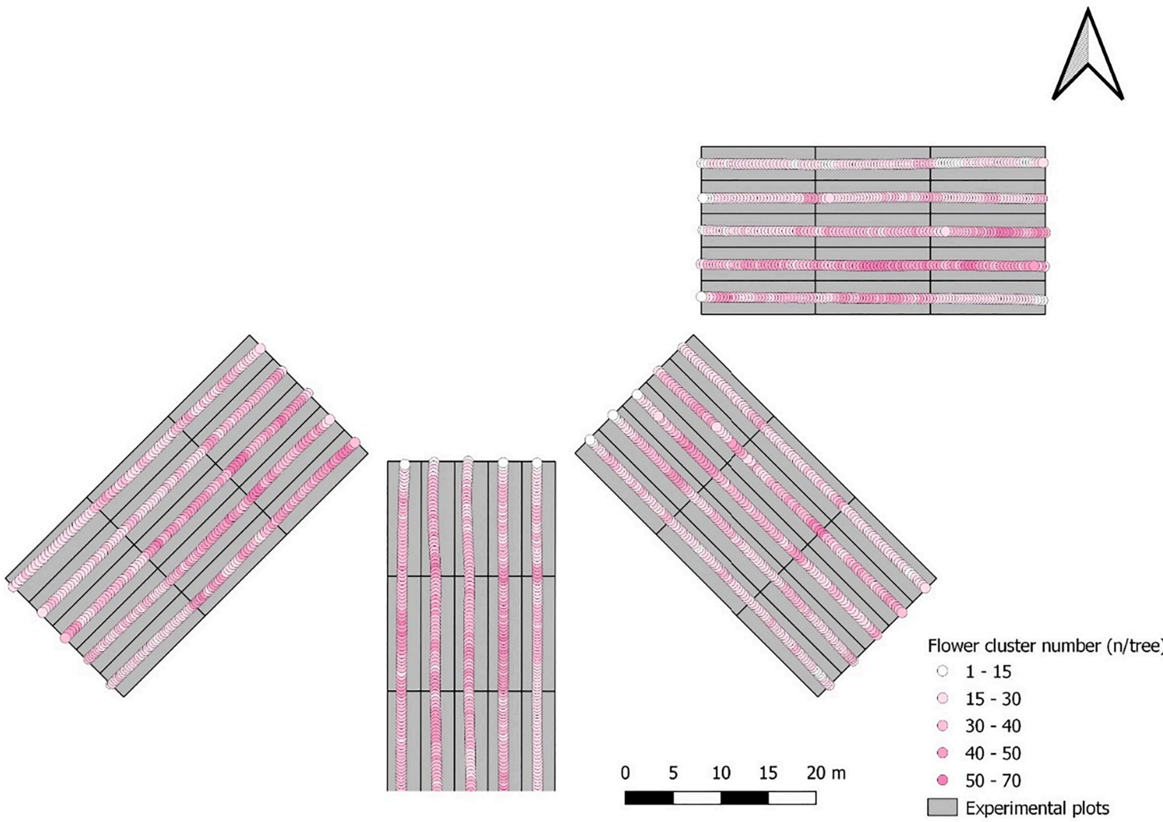

Fig. 4. Heatmap of calibrated flower cluster number in the sixty experimental plots of 'ANABP-01′ apples. Data collected at full bloom.

the LiDAR-generated points in the scanned transect, excluding the trunk; canopy density represents the ratio between the number of light beams generated by the LiDAR that bounces back to the light source and the total number of emitted light beams within the canopy area; and canopy cross-sectional leaf area (CSLA, m^2) can be calculated as the product of canopy area and canopy density and is equivalent to the area of the points (comparable to leaves) within the canopy area polygon in the scanned transect. Green Atlas is rapidly expanding *Cartographer*'s capability and aims to achieve good predictions of flower cluster number, fruit size and fruit surface colour. The use of artificial intelligence to estimate fruit number and potentially fruit size and weight, will expand and improve the ability of orchardists to predict yield. According to Anderson et al. (2021), a generally accepted error of yield estimations is 5 – 10 %, but this figure differs across fruit industries.

The overall objective of this study was to evaluate *Cartographer* as a rapid orchard assessment tool to determine several crop parameters that are typically measured manually and then use calibrated *Cartographer* data to explore the effects of agronomic treatments on productivity of 'ANABP-01′ apple trees. Specifically, this work aimed to: (i) establish relationships between manual measurements (i.e. observed variables) and *Cartographer* scans (i.e. detected variables) of flower cluster number, fruit number, yield and tree height; (ii) establish the relationship of LiDAR-obtained tree geometry parameters (i.e. canopy area, canopy density and CSLA) with light interception; (iii) determine the effects of light interception on flower and fruit density and yield; and (iv) estimate the effects of rootstock and row orientation on flower cluster number, fruit number, yield, tree height and tree geometry.

2. Materials and methods

2.1. Experimental site and apple cultivar

The study was conducted in the Sundial orchard at the Tatura SmartFarm, Victoria, Australia during 2020 – 2021. The Sundial orchard is a high-density (HD, ~ 2857 trees / ha) circular orchard of approximately 1.3 ha. 'ANABP-01′ (marketed as Bravo™) apple trees were planted in a semicircle of the orchard following four different row orientations (N – S, NE – SW, E – W and SE – NW). Trees were grafted onto three different rootstocks [Bud.9, M9 (T337) and M26], planted in 2018 at 1 m tree spacing and 3.5 m row spacing, and trained to Spindles on a vertical trellis. There was a total of twenty rows, with five rows per row orientation, and 60 experimental plots. Each experimental plot was composed of eleven 'ANABP-01′ trees and one polleniser ('Granny Smith'). The experimental design was completely randomised with rootstock and row orientation as factors (Fig. 1). The experiment was conducted on trees in their 3rd leaf.

'ANABP-01′ originated from a cross-pollination between 'Cripps Red' and 'Royal Gala'. The cultivar was bred by the Department of Agriculture and Food, State of Western Australia. Fruit has dark purple colouration and consistent cropping characteristics (Cripps, 2016).

2.2. Orchard scans with the mobile platform

Orchard scans were undertaken with a commercial orchard scanning product called *Cartographer*, commercialised by the Australian company Green Atlas. The system combines LiDAR, GPS and cameras with state-of-the-art machine vision algorithms, including point cloud processing and convolutional neural networks (CNN) to map orchards with precision. A smartphone interface was used to control logging and enter experimental metadata to aid retrospective identification of scan locations and note relevant scan or plot issues. *Cartographer* was driven at a constant speed of approximately 7 – 8 km/h for calibration scans, and at 20 km / h when mapping orchard blocks. Logging was switched on a few metres prior to the start of the measurement section and off a few metres past the end of the measurement section. Images were logged at a rate of 5 images / second.

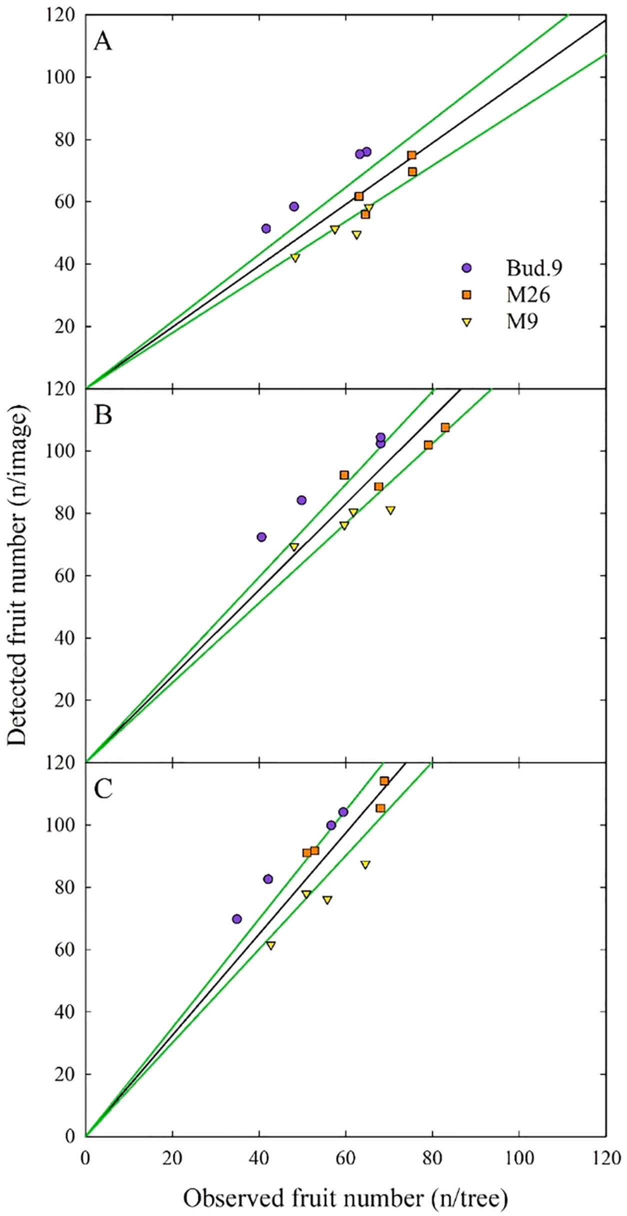

Fig. 5. Detected against observed fruit number in 'ANABP-01' apple trees grafted on three rootstocks at (A) 44; (B) 102; and (C) 154 days after full bloom. Black and green lines represent linear regression fits and 95% confidence interval bands, respectively. (A) y = 0.99 (0.04) ×, *RMSE* = 8 fruit / image; (B) y = 1.38 (0.02) ×, *RMSE* = 10 fruit / image; (C) y = 1.62 (0.02) ×, *RMSE* = 10 fruit / image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Short mobile scans were conducted to generate counts of flower clusters and fruit per image. Experimental plots were scanned on both sides. These scans were collected in sections of the rows corresponding to the experimental plots. Detections were calibrated against manual measurements using linear regression procedures with intercept set to 0. A calibration factor (i.e. the inverse of the slope of a linear regression between detections and manual measurements) was used to adjust detected counts generated by *Cartographer*.

Continuous mobile scans were collected in the entire orchard block to obtain uncalibrated predictions of flower and fruit counts and to assess tree geometry (i.e. tree height, canopy area, canopy density and CSLA). Flower clusters and fruit count detections obtained from continuous mobile scans of all the measurement rows in the Sundial orchard were reprocessed using the calibration factor. Tree geometry data was used with no additional calibration.

Table 1
Calibration factors and root mean square errors (*RMSE*) of fruit number detected with *Cartographer* in 'ANABP-01′ apple trees on three rootstocks and at three scan dates.

Days after full bloom	Rootstock	Calibration factor	*RMSE* (fruit n / image)
44	Bud.9	0.837	1
	M9	1.059	2
	M26	1.161	3
102	Bud.9	0.632	6
	M9	0.746	7
	M26	0.787	6
154	Bud.9	0.548	5
	M9	0.602	5
	M26	0.708	4

Data at a plot level was extracted by intersecting the points generated by *Cartographer* with plot polygons generated with QGIS (v.3.10, QGIS Development Team, 2021). The geographic precision of plot extraction was improved by real-time kinematic (RTK) positioning adjustments and was cross-checked on the position of posts — to mark the beginning and end of a plot — in the RGB images collected by *Cartographer*.

Stationary scans were conducted to compare detected tree height with manual measurements to simplify extraction of individual tree height data.

2.3. Predictions of tree parameters

2.3.1. Flower cluster number

Cartographer was used to collect short mobile scans in six experimental plots at 50 % bloom (1 October 2020) for the calibration of flower cluster counts. On the same day, manual measurements of flower clusters were obtained by counting clusters on all the trees in the same plots. Clusters at phenological stages pink balloon to open cluster (i.e. at least one fully open flower in the cluster) were counted to obtain the total cluster count per plot. Manual measures of clusters per tree and detected counts per image were associated by linear regression analysis with intercept = 0 to determine the calibration factor. Flower counts were not validated due to the absence of an independent separate validation dataset.

All the apple plots in the Sundial orchard were then scanned continuously at full bloom to obtain uncalibrated numbers of flower clusters that were subsequently adjusted using the calibration factor.

2.3.2. Fruit number and yield

Fruit number was determined in twelve experimental plots in the Sundial orchard. Measurements were done at three stages to target different fruit size. Plots were scanned at 44, 102 and 154 days after full bloom (DAFB). Short mobile scans were conducted as done for flower clusters. Fruit was manually counted on each day the scans were obtained. Linear regression models of detected counts per image against the manual counts were used to determine calibration factors to correct fruit number predictions at the three observation dates. The robustness of the calibrated model of fruit count was validated against a dataset of counts per plot obtained with a commercial grader (Compac InVision 9000, Compac Sorting Equipment Ltd, Australia) at harvest on 36 experimental plots.

Cartographer predictions of yield per plot and per tree were obtained by multiplying fruit number by average fruit weight per rootstock. Fruit weight (FW) was calculated based on the exponential relationship between fruit weight and fruit equatorial diameter (FD) determined from data collected during the growing season (Fig. 2). FD was determined with a digital calliper on 108 fruits, 36 per rootstock, in the week preceding harvest. Yield predictions were validated against the yield obtained with the commercial grader at harvest on 36 experimental plots.

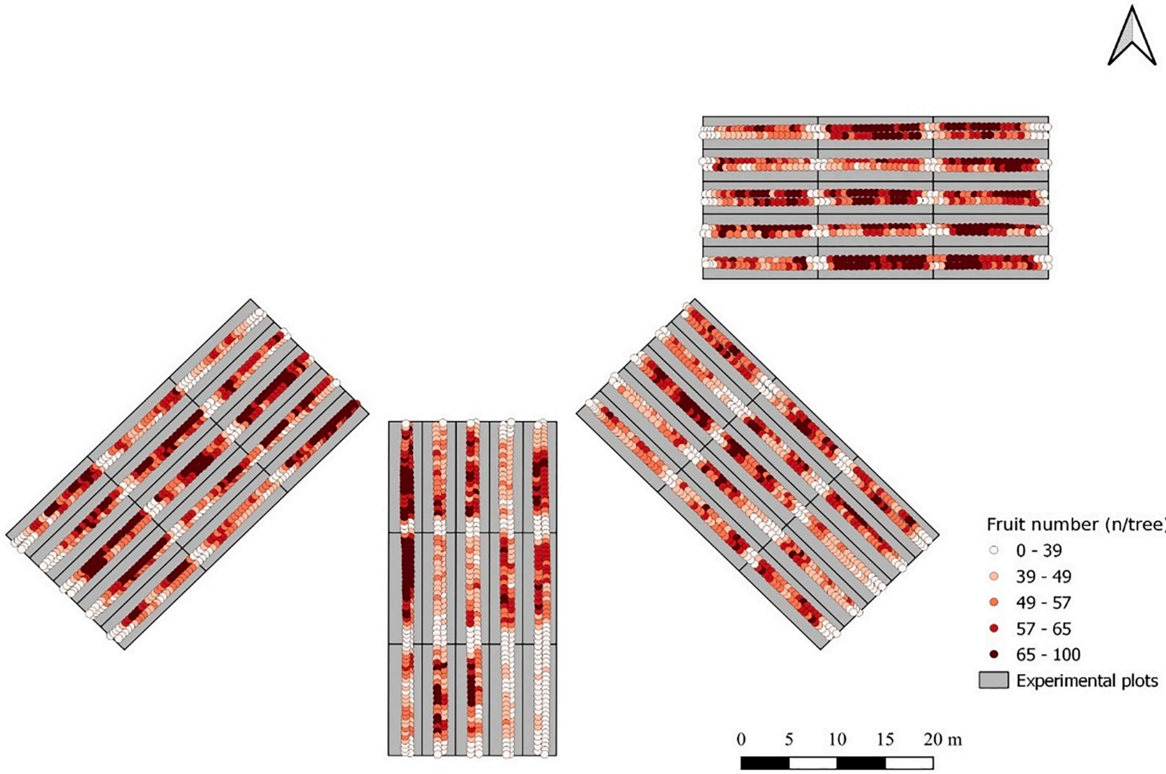

Fig. 6. Heatmap of calibrated fruit number in the sixty experimental plots of 'ANABP-01′ apples. Data was collected a week prior to harvest (154 days after full bloom).

2.3.3. Tree height

Cartographer stationary scans were collected on the central trees of the 36 plots (Fig. 1) within the Sundial orchard at 45, 101 and 154 DAFB. Stationary images were collected while cameras were pointed to the tree. Stationary scans were used to detect uncalibrated tree height. Predictions were obtained adjusting detections by subtracting the error of the linear model, that represented the bias in the detection of ground level. Tree height was predicted using LiDAR images, whereas manual measurements were collected with a measuring stick. Normally, *Cartographer* measures canopy geometry continuously without stopping; in our case, stationary scans were used to guarantee the correspondence of the measurement to the same tree measured manually in the field.

2.4. Relationship between light interception and tree geometry

Canopy light interception was compared to LiDAR-obtained tree geometry parameters canopy area, canopy density and CSLA. Light interception was expressed in terms of effective area of shade (EAS, Goodwin et al., 2006) — the mean of fractional photosynthetically active radiation (PAR) interception over the tree planting square (tree × row spacing) measured at three times (solar noon, solar noon − 3.5 h and solar noon + 3.5 h) on a clear sky day. PAR was measured using a light trolley (Tranzflo, Palmerston North, New Zealand). The light trolley consisted of 24 PAR sensors at 0.125 m intervals along a 3 m bar, 0.4 m above ground level on a wheeled base. An on-board data logger (CR850, Campbell Scientific, Garbutt, Au) recorded measurements at 1 s intervals. Measurements of transmitted PAR (PARt) were made over the planting square of the central trees in each plot. The light trolley sensors were held horizontally below the canopy, perpendicular to the row direction, and moved at a slow walking speed. Unobstructed incoming PAR (PARi) was measured at 1.5 m above ground level in an open area. Measurements with the light trolley were carried out at two dates (44 and 102 DAFB) in the 60 experimental plots.

Cartographer scans were collected on the 60 experimental plots within the Sundial orchard at three dates (44, 102 and 154 DAFB). Data of canopy area, canopy density and CSLA per plot were extracted using QGIS by intersecting a polygon layer containing the 60 experimental plots and the points generated by *Cartographer*. The relationships of canopy area, canopy density and CSLA against EAS were modelled by linear regression analysis using medians per plot at 44 and 102 DAFB. Models were separately obtained for the two dates and the line intercepts and slopes were compared to determine whether time of the year affected the relationship between tree geometry and EAS. The aim was to find the tree geometry parameter that had the most robust relationship with EAS regardless of date of measurement. After determining which tree geometry was the best predictor of light interception, significant effects of rootstock and row orientation on the relationships between tree geometry and EAS were investigated by comparing slopes and intercepts between treatments.

2.5. Effects of CSLA on flower clusters, fruit number and yield

Flower cluster, fruit number and yield predictions in the sixty experimental plots were related to CSLA measured at 154 DAFB to determine whether there was an effect of canopy size on productivity.

2.6. Effects of row orientation and rootstock

The effects of row orientation and rootstock on predicted flower cluster number, fruit number, yield, tree height, canopy area, canopy density, CSLA and EAS a week prior to harvest were analysed to summarise results for the seasonal data.

2.7. Statistical analysis

Model prediction errors of both calibrations and validations were based on root mean square errors (*RMSE*) of the linear regressions. The robustness of validation models was assessed using the Lin's

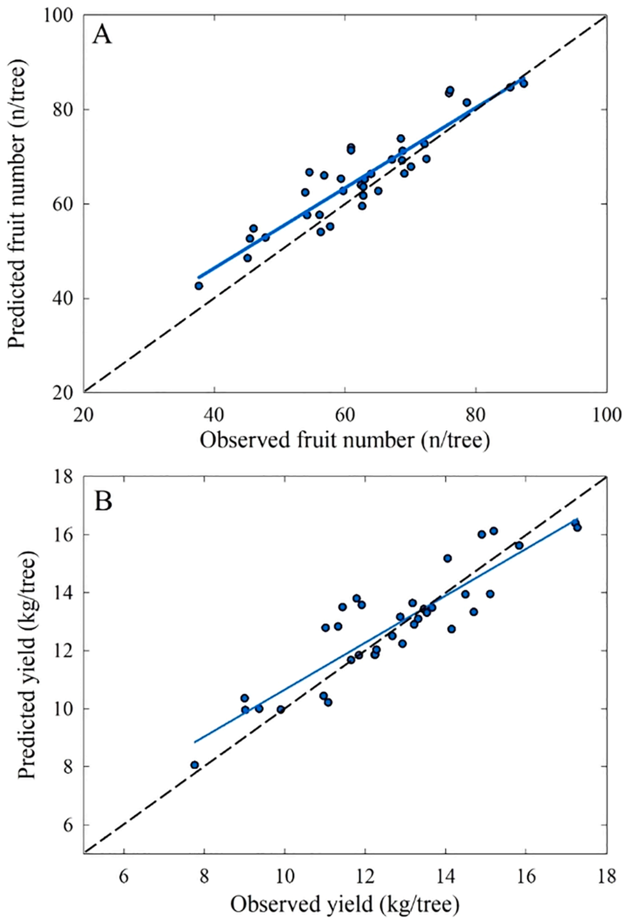

Fig. 7. Validations of *Cartographer* predictions of (A) fruit number; and (B) yield at 154 DAFB (i.e. a week prior to harvest) against fruit number and yield measured by a commercial grader at harvest. Blue lines are the linear regression fits; grey dashed lines represent the y = x fit. (A) y = 12.4 (4.0) + 0.85 (0.06) × [r_c = 0.88; *RMSE* = 5 fruit / tree]; and (B) y = 2.57 (0.86) + 0.81 (0.07) × [r_c = 0.89; *RMSE* = 1 kg / tree]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

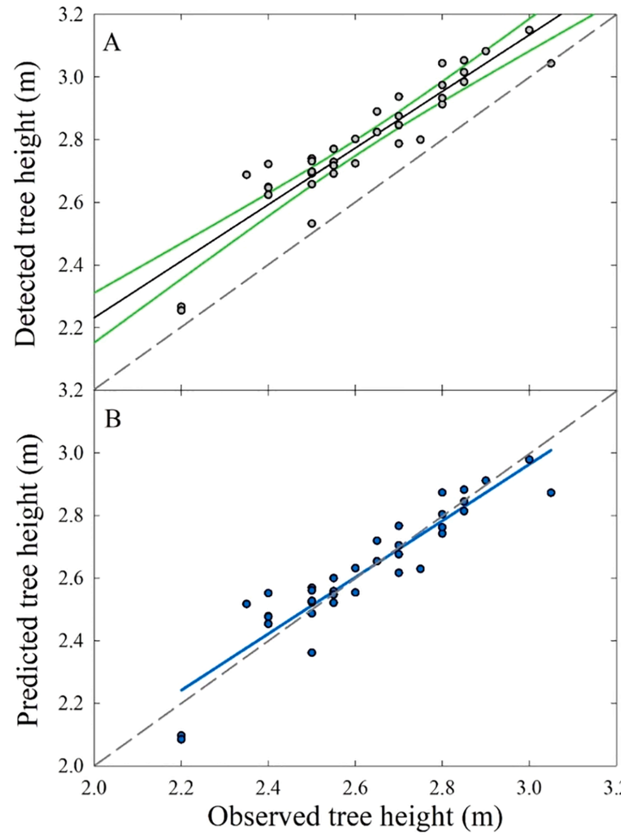

Fig. 8. Detected (uncalibrated) and predicted (calibrated) tree heights against observed tree height in 'ANABP-01′ apple trees. (A) uncalibrated data [y = 0.43 + 0.90 x]; and (B) data calibrated subtracting 0.171 m from predictions (i.e. error in ground height) [y = 0.26 + 0.90 ×, r_c = 0.93]. Black line is the uncalibrated linear regression fit; green lines show the 95% confidence interval bands; blue line is the calibrated linear regression fit; grey dashed lines: y = x fit. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

concordance correlation coefficient (r_c), a measure of both precision and accuracy that provides a value from 0 (no concordance) to 1 (perfect concordance) that assesses the divergence of predicted data from the line of perfect concordance with observations (i.e. the line at 45 degrees on a square scatter plot, where y = x) (Lin, 1989).

The effects of light interception on fruit number and yield were estimated by correlation analysis and assessed with the Pearson's correlation coefficient (r). The effects of row orientation and rootstock were tested using a two-way analysis of variance (ANOVA) and significant differences ($p < 0.05$) were separated by Tukey's Honestly Significant difference (HSD). Regression analyses, ANOVA, post hoc tests and the calculation of Lin's concordance correlation coefficients for validation models (r_c) were carried out using R (v. 4.0.2, R Core Team. R: A (2018), R Core Team, 2021) and its packages "Userfriendliscience" (Peters, 2018) and "DescTools" (Signorell et al., 2021). Graphs were generated using SigmaPlot 12.5 (Systat software Inc., Chicago, IL, USA) and heatmaps were produced with QGIS.

3. Results

3.1. Predictions of tree parameters

3.1.1. Flower cluster number

Fig. 3 shows the relationship between *Cartographer* detected counts of flower clusters per image and observed values of flower clusters per tree at 50 % bloom. The number of detected flower clusters was tightly associated with observations and the linear regression model returned a low prediction error (*RMSE* = 5 clusters per image). Calibrated flower cluster counts are presented in a heatmap overlayed to the experimental plots (Fig. 4).

3.1.2. Fruit number and yield

The relationship between detected fruit number per image and observed fruit number per tree demonstrated similar calibration robustness in the three dates when measurements were carried out, although the slopes of the lines were visibly different (Fig. 5). Prediction errors were similar over time (i.e. *RMSE* = 8 – 10n / image).

The scatterplots in Fig. 5 show that the three rootstocks appeared to form different groups (Fig. 5). Therefore, calibration factors (i.e. slopes) and prediction errors were recalculated independently for each rootstock and date (Table 1). The *RMSE* values decreased when rootstocks were separated (Table 1) compared to when they were pooled together (Fig. 5). Fig. 6 shows an example of a fruit number heatmap of the Sundial orchard (154 DAFB) calibrated using the calibration factors in Table 1. Validations of fruit number and yield estimates obtained with *Cartographer* (Fig. 7) against data from the commercial grader highlighted strong reliability of the predictions at 154 DAFB ($r_c > 0.85$) and small prediction errors (*RMSE* = 5 fruit / tree and *RMSE* = 1 kg / tree, respectively).

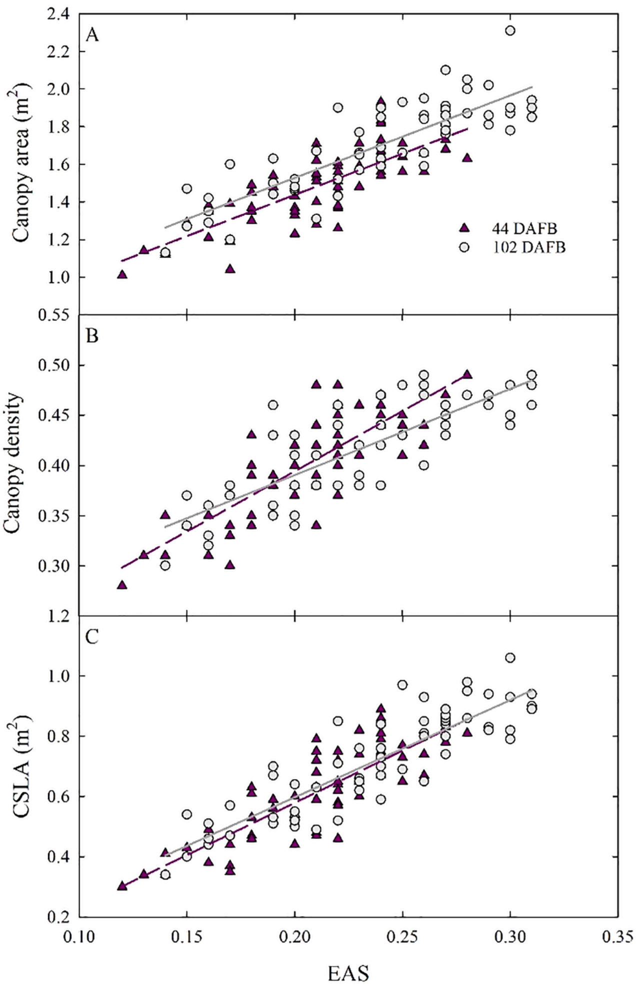

Fig. 9. Scatterplots and linear regression fits of (A) canopy area; (B) canopy density; and (C) canopy cross-sectional leaf area (CSLA) measured by *Cartographer* against effective area of shade (EAS) measured by a light trolley on two dates (in days after full bloom, DAFB). Points represent medians of 60 experimental plots. Linear models reported in Table 2.

Table 2
Linear regression models of tree geometry parameters against effective area of shade at two dates (in days after full bloom, DAFB). Coefficients of determination (R^2) and analysis of variance *p*-values for intercept and slope comparisons between dates. Standard errors of slope and intercept values are reported in brackets.

Tree geometry	DAFB	Equation	R^2	Intercept *p*-value	Slope *p*-value
Canopy area	44	y = 0.56 (0.09) + 4.38 (0.44)	0.71	0.491	0.990
	102	y = 0.65 (0.09) + 4.38 (0.37)	0.62		
Canopy density	44	y = 0.15 (0.02) + 1.20 (0.10)	0.70	0.027	0.010
	102	y = 0.22 (0.02) + 0.86 (0.08)	0.66		
Canopy cross-sectional leaf area	44	y = − 0.11 (0.06) + 3.46 (0.29)	0.70	0.439	0.334
	102	y = − 0.05 (0.06) + 3.23 (0.24)	0.76		

Table 3
Linear regression models of canopy cross-sectional leaf area (CSLA) against effective area of shade in four row orientations and three rootstocks. Analysis of variance *p*-values for intercept and slope comparisons. Standard errors of slope and intercept values are reported in brackets.

Factor	Level	Equation	Intercept *p*-value	Slope *p*-value
Row orientation	E – W	y = − 0.09 (0.08) + 3.66 (0.34) x	0.221	0.478
	NE – SW	y = − 0.29 (0.11) + 3.99 (0.48) x		
	N – S	y = − 0.12 (0.09) + 3.55 (0.43) x		
	NW – SE	y = − 0.08 (0.11) + 3.23 (0.50) x		
Rootstock	Bud.9	y = 0.11 (0.07) + 2.04 (0.38) x	0.258	0.129
	M26	y = 0.02 (0.11) + 3.01 (0.51) x		
	M9	y = 0.21 (0.11) + 2.26 (0.52) x		

3.1.3. Tree height

Tree height detections showed a good association with manual measurements (Fig. 8A). When the slope was set to 1, the model with the highest coefficient of determination ($R^2 = 0.858$) had an intercept of 0.171 m that represented an error in ground height detected by *Cartographer*. The intercept value was subtracted from detected height to calibrate detections and the linear association became much closer to the y = x line (Fig. 8B). Good reliability of the prediction was reflected in a resulting r_c of 0.93.

3.2. Relationship between light interception and tree geometry

Fig. 9 shows the linear relationships of canopy area, canopy density and CSLA with EAS at 44 and 102 DAFB. The three geometry parameters had a positive linear association with EAS — i.e. more light was intercepted at increasing canopy area, density and CSLA. When the two dates were compared, the linear regression models were found to be significantly different for canopy density (i.e. intercept and slope *p*-values in Table 2). The slopes and intercepts of canopy area and CSLA against EAS relationships were not significantly different between dates (Table 2). The latter generated the most robust association with EAS, regardless of date of measurement ($R^2 \geq 0.70$). The model for the estimation of EAS from CSLA is reported in Eq (1).

$$EAS = 0.07\ (0.01) + 0.23\ (0.01)\ CSLA\ \left[R^2 = 0.76,\ RMSE = 0.03\right] \quad (1)$$

No significant effects of row orientation and rootstock were observed on the relationship between CSLA vs EAS, as slopes and intercepts of the models were not significantly different (Table 3). An orchard heatmap of CSLA is presented in Fig. 10.

3.3. Effects of CSLA on flower number, fruit number and yield

Correlation analysis showed that in the CSLA range of 0.3 – 1.1 m^2, increasing canopy size was associated with a higher number of flower clusters, and an increase in fruit number and final yield (Fig. 11).

3.4. Effects of row orientation and rootstock

Row orientation and rootstock had a significant effect on the parameters generated by *Cartographer* but no significant interactions between the two factors were observed (Table 4).

Trees planted in E – W and NE – SW rows bore more flower clusters and fruit and had higher yield compared to the other row orientations. E – W trees were taller and had the largest canopy area and CSLA. Trees in NE – SW rows were denser and intercepted the highest proportion of

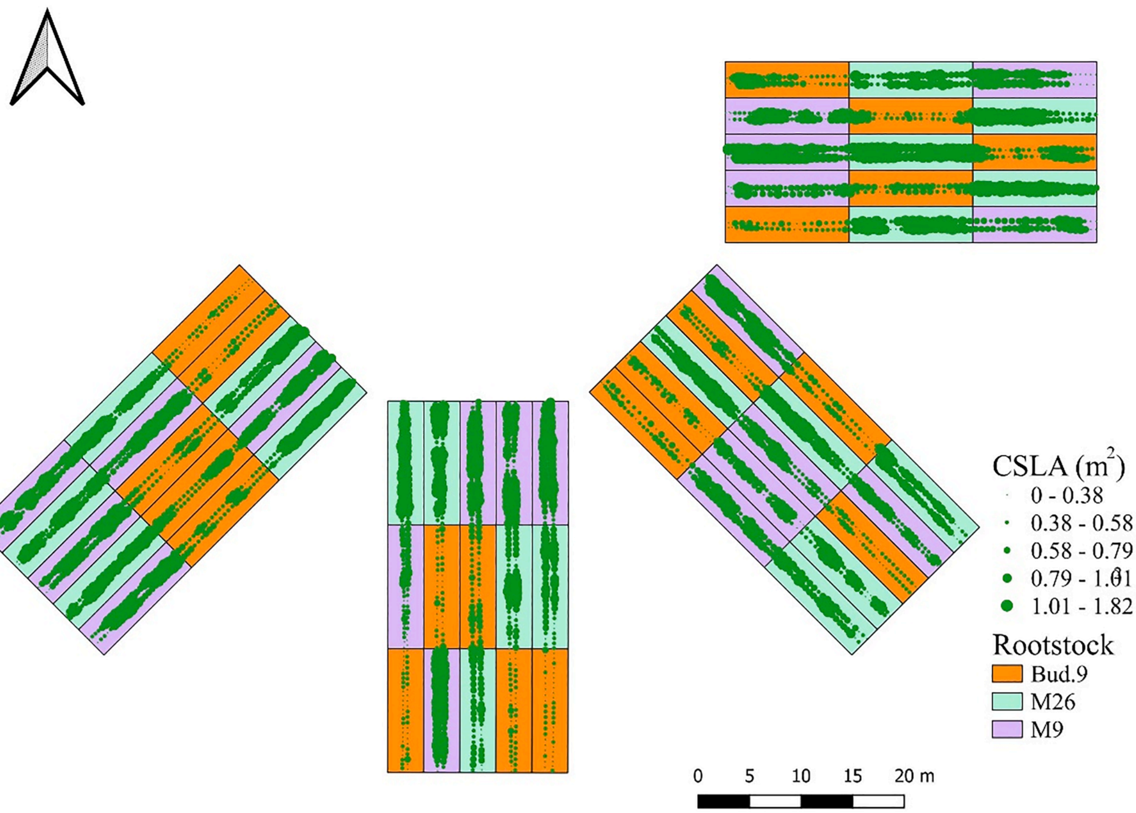

Fig. 10. Heatmap of canopy cross-sectional leaf area (CSLA) measured by *Cartographer* in the sixty experimental plots of 'ANABP-01′ apples at 154 days after full bloom.

light (i.e. EAS) although canopy area and CSLA were smaller than E – W trees. Differences in terms of canopy density and CSLA between row orientations were not as marked as for the other parameters, as their p approached $\alpha = 0.050$ ($p = 0.042$, 0.030 and 0.050, respectively) and their effect size (η^2) was ≤ 0.05 (Table 4).

Among the rootstocks, M26 trees bore more flower clusters and fruit, and had the highest yield and tree height (Table 4). M9 trees had similar canopy area, canopy density, CSLA to M26 trees. Bud.9 trees were significantly smaller in terms of canopy area and CSLA, had lower density and intercepted less light than trees grafted on M9 and M26.

4. Discussion

Overall, *Cartographer* provided accurate predictions of several important crop parameters in 'ANABP-01′ apples. Its use can be beneficial both for growers and scientists to collect data for multiple crop properties at high spatial resolution and replace labour-intensive operations.

The association between *Cartographer* detections of flower cluster numbers and observations was very tight and produced a small error (Fig. 3). The overestimation of flower cluster number detected by *Cartographer* was due to a combination of factors, including (i) detections from next row (false positives), (ii) detections from adjacent trees within the row (+false positives), (iii) detections of flowers within the cluster as individual clusters (+false positives), and (iv) missed detections (false negatives). Uncalibrated orchard maps can be produced to display areas or rows in the block where flower clusters are denser than others, to support precise chemical (e.g. variable rate sprayers) or mechanical thinning. A first calibration step is needed if absolute flower cluster numbers are to be determined and displayed in spatial maps (Fig. 4). Green Atlas provides calibrated or uncalibrated heatmaps to growers as part of their commercial service.

Fruit number calibrations had an error ranging from 8 to 10 fruit per image at three different dates during summer, when data from trees grafted on Bud.9, M9 and M26 were pooled together (Fig. 5). When rootstocks were separated, the error was significantly reduced (Table 1); thus, carrying out independent calibrations when within-block conditions change (e.g. rootstock, tree spacing, row spacing, tree architecture) helps minimise detection errors. Bud.9 trees had consistently the smallest calibration factor at the three measurement dates (Table 1). The predictive algorithm likely detected more false positive fruit in Bud.9 compared to M9 and M26. This may have been caused by increased fruit detections from trees in the next row, as Bud.9 trees had significantly smaller canopy area, canopy density and CSLA than M9 and M26 trees (Table 3).

Validations of fruit number and yield predictions a week prior to harvest highlighted good concordance between observations and predictions (Fig. 7). Block estimates of fruit number and yield were considered very accurate, as prediction errors were 4.6 and 1.2 %, respectively. These results can be considered very good if compared to the generally accepted yield prediction error of 5 – 10 % reported by Anderson et al. (2021). Yield predictions presented in this study represent one of the first attempts to predict this parameter in apples. The *RMSE* of yield prediction in this study (1 kg / tree) was lower than the one obtained with a back propagation neural network in 'Gala' apples (Cheng et al., 2017), who found an error of 2.5 – 2.6 kg / tree, although the trees used by Cheng et al. (2017) had a slightly higher yield (18 kg / tree) than the ones used in this study (13 kg / tree).

Using the methodology presented in this study, yield could be predicted with relatively low error any time after fruit thinning or natural fruit drop. Early predictions represent valuable informative tools to estimate revenues and support decision management in the logistics and post-harvest handling of the crop. Like for flowers, calibration of the fruit number estimations is only needed if absolute fruit number and yield estimates are needed, and in this case, spatial maps displaying absolute values can easily be obtained (Fig. 6). In several circumstances,

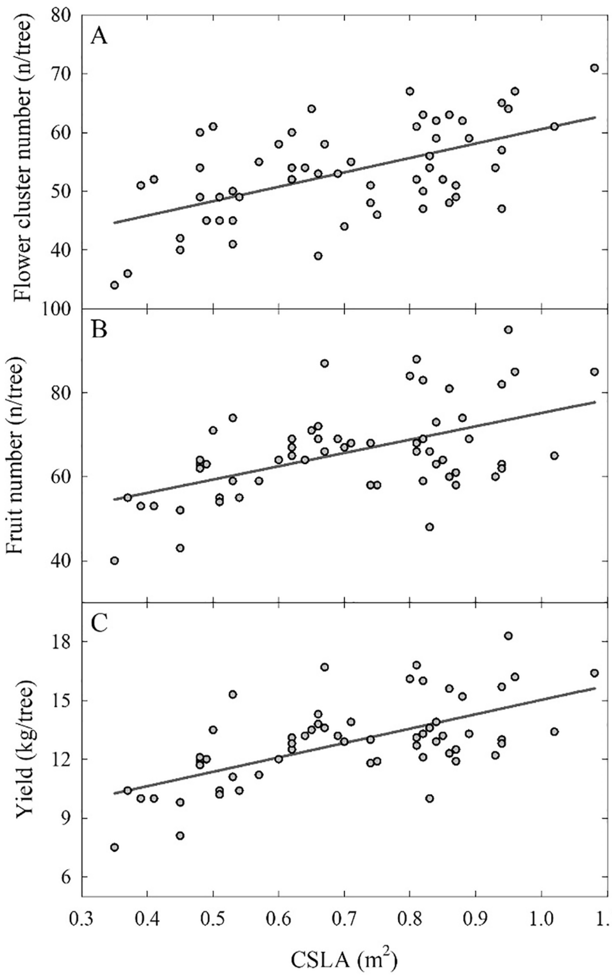

Fig. 11. Scatterplots of (A) flower clusters, (B) fruit number and (C) yield plotted against canopy cross-sectional leaf area (CSLA) in 'ANABP-01′ apples. Correlation analyses: (A) Pearson's $r = 0.562$, $p < 0.001$; (B) $r = 0.531$, $p < 0.001$; (C) $r = 0.631$, $p < 0.001$. Data from sixty experimental plots.

an uncalibrated, relative fruit number heatmap of the orchard is sufficient to support thinning management decisions (e.g. management of labour for thinning operations).

Tree height predictions needed a preliminary calibration for the ground height (Fig. 8A). After calibration, tree height predictions were considered accurate and in line with manual observations (Fig. 8B). A reliable prediction of tree height is important to support pruning management and automation. Furthermore, in high-density orchards trees are often trained and managed to have an optimal height, so that the amount of light intercepted by trees is optimised and excessive shading on next-row trees can be avoided.

Tree geometry parameters generated by the LiDAR on *Cartographer* were related to light interception measurements. Specifically, CSLA had the most robust and stable relationship with EAS ($R^2 \geq 0.70$; Table 2) and this relationship was not affected by row orientation and rootstock (Table 3). Row orientation is expected to influence the amount of intercepted light by trees with similar canopy size based on empirical models (Palmer, 1989; Trentacoste et al., 2015). In our study, trees were young, and canopies not fully developed, achieving maximum EAS < 0.35 in the Sundial orchard experimental plots. This might have influenced the lack of a significant effect of row orientation on the relationship between canopy size and intercepted light. It is important to reassess row orientation effects when trees will have fully developed canopies that have EAS > 0.35. CSLA, canopy area and density can be used per se to assess canopy size and locate tree architecture irregularities or gaps in an orchard.

Georeferenced orchard heatmaps of tree geometry (e.g. CSLA, Fig. 10) can support precise and targeted management of pruning, fertilisation and replanting. The relationships between tree geometry parameters measured by *Cartographer*, particularly CSLA, and EAS will enable greatly increased assessment of canopy light interception both in research and commercial orchards. Canopy light interception is a major determinant of crop water use (Allen et al., 1998, Goodwin et al., 2006) and can be used to improve irrigation scheduling either by matching irrigation supply to mean crop water requirement at the block scale or by implementing irrigation management units based on EAS (McClymont et al., 2011, 2012). However, orchardists have found traditional methods of assessing EAS to be difficult. A tool such as *Cartographer* would greatly increase the opportunity to improve irrigation management based on EAS.

Flower cluster number, fruit number and yield were positively affected by CSLA (Fig. 11). This might have been due to an indirect beneficial effect of increased light interception,rather than to CSLA per se, in line with previous findings on 'Empire' (Robinson and Lakso, 1989; Wünsche et al., 1996; Wünsche and Lakso, 2000) and 'Elstar' (Wagenmakers and Callesen, 1995) apples. Relationships of flower cluster number, fruit number and yield with CSLA need to be investigated at CSLA > 1.10 m^2 and EAS > 0.35 to determine the optimal level of canopy size and light interception to achieve the highest productivity. In this study, yield was obtained by multiplying fruit number by average fruit weight. Fruit number is the primary determinant of yield and minor errors in fruit weight will still allow for relatively accurate yield predictions. However, the ability to estimate fruit diameter in addition to predicting fruit number could further improve yield predictions.

Row orientation significantly affected the number of flower clusters, fruit number and yield. E – W trees produced more fruit, had higher yield and canopy area (Table 3). Similarly, higher yields in E – W rows with respect to N – S rows were obtained by Devyatov and Gorny (1978) in apple and by Gómez-del-Campo et al. (2009) in olive. On the other hand, Christensen (1979) and Middleton et al. (1982) found higher productive performance in N – S than in E – W row orientations in apple. At a similar latitude to the one in this study, Hunter et al. (2017) observed slightly higher yields in N – S vineyard rows, and slightly higher berry mass in E – W orientations. Altered microclimate and differences in diurnal patterns of light interception occur in response to different row orientations (Palmer 1989, Hunter et al., 2017). The nature of these changes and the subsequent impact on tree physiological status, growth, yield performance and fruit quality is influenced by latitude, local environmental conditions, macroclimate, training system and genotype. Future investigation of the drivers of row orientation effects at this site is required and will add to the ability to model potential effects of row orientation in different environments or on different crops.

Rootstock effects appeared strong and M26 trees were generally larger and yielded more flower clusters and fruit (Table 3), whereas Bud.9 trees were more compact, intercepted less light and had a smaller yield than M9 and M26. In a six-year study, Univer et al. (2017) observed significantly higher yields and vigour in 'Krista' apples grafted on M26 over Bud.9; M9 was not part of their study. Overall, in our study, rootstock differences were more marked than row orientation differences, although significant in both cases, as suggested by the generally higher ANOVA's effect sizes in the former (Table 3). The only exceptions occurred for flower cluster number and tree height, which seemed to be similarly affected by row orientation and rootstock.

Current research is assessing the ability to estimate fruit size and colour to produce orchard maps that will further support orchard management. Accurate spatial distributions of fruit diameter and colour will potentially provide zonal information on fruit quality and will support orchard management practices such as thinning, use of reflective films and selective harvest.

Table 4
Effects of row orientation, rootstock, and their interaction, on the variables obtained with *Cartographer*. ANOVA's *p*-values, effect size (η2) and Tukey's Honestly Significant Difference (HSD) reported. One, two and three asterisks represent < 0.05, 0.01 and 0.001 significance levels, respectively; n.s.: non-significant ($p > 0.05$).

Effect	Level	FCN[q]	FN[r]	YI[s]	TH[t]	CA[u]	CD[v]	CSLA[w]	EAS[x]
RO[y]	E – W	57	72	14.0	2.71	1.82	0.41	0.76	0.24
	NE – SW	57	67	13.1	2.61	1.66	0.43	0.73	0.26
	N – S	49	61	11.8	2.61	1.59	0.41	0.66	0.22
	NW – SE	50	63	12.3	2.65	1.62	0.41	0.67	0.23
	p	***	***	***	**	***	*	*	**
	η^2	0.19	0.15	0.15	0.19	0.13	0.04	0.05	0.10
	HSD	6	6	1.2	0.07	0.15	0.02	0.10	0.03
RS[z]	Bud.9	50	58	10.9	2.62	1.43	0.35	0.50	0.19
	M26	58	76	14.7	2.69	1.79	0.45	0.80	0.26
	M9	52	63	12.9	2.62	1.81	0.44	0.81	0.26
	p	***	***	***	**	***	***	***	***
	η^2	0.19	0.49	0.52	0.12	0.51	0.70	0.65	0.56
	HSD	5	5	1.0	0.06	0.12	0.02	0.08	0.02
RO × RS	*p*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	η^2	0.05	0.07	0.06	0.13	0.04	0.04	0.03	0.01
	HSD	–	–	–	–	–	–	–	–

[q] Flower cluster number (n / tree).
[r] Fruit number (n / tree).
[s] Yield (kg / tree).
[t] Tree height (m).
[u] Canopy area (m^2).
[v] Canopy density.
[w] Canopy cross-sectional leaf area (m^2).
[x] Effective area of shade.
[y] Row orientation.
[z] Rootstock.

5. Conclusions

In summary, in this study we calibrated and validated a mobile platform for the prediction of several crop parameters in 'ANABP-01′ apples. Combining predictions of several important parameters in one single platform opens the door to various possible uses of this technology. With the current push of AgTech into business models in the apple industry, and in other fruit industries, the availability of technology that serves multiple purposes is of pivotal importance to reduce substantial production costs such as labour and to ease growers' technology uptake in their business models and investment plans. Overall, *Cartographer* demonstrated to be a valid tool to combine predictions of several important fruit crop parameters (i.e. flower cluster number, fruit number, yield, tree size and geometry) in one single platform and its use can be beneficial both for growers and scientists to collect large amount of data and replace labour-intensive operations.

Funding

This study was a component of the apple and pear industry's PIPS3 (Productivity, Irrigation, Pests and Soils) program of research and development funded by Hort Innovation, using the Hort Innovation Apple and Pear research and development levy, contributions from the Australian Government and co-investment from Agriculture Victoria. Hort Innovation is the grower-owned, not-for-profit research and development corporation for Australian horticulture.

CRediT authorship contribution statement

Alessio Scalisi: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Visualization. **Lexie McClymont:** Conceptualization, Methodology, Writing – review & editing. **James Underwood:** Software, Resources, Funding acquisition. **Peter Morton:** Software, Resources. **Steve Scheding:** Software, Resources, Funding acquisition. **Ian Goodwin:** Conceptualization, Methodology, Resources, Writing – review & editing, Project administration, Funding acquisition.

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.

Acknowledgements

The technical support and assistance of Dave Haberfield, Michael Halverson, and Laura Phillips is gratefully acknowledged.

References

Allen, R., Pereira, L.S., Raes, D., Smith, M. 1998. Crop evapotranspiration. Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. FAO, Rome, Italy.
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), 1409. https://doi.org/10.3390/agronomy11071409.
Bargoti, S., Underwood, J., 2017a. Deep fruit detection in orchards, in: Proceedings - IEEE International Conference on Robotics and Automation. Institute of Electrical and Electronics Engineers Inc., pp. 3626–3633. 10.1109/ICRA.2017.7989417.
Bargoti, S., Underwood, J.P., 2017. Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards. J. F. Robot. 34, 1039–1060. https://doi.org/10.1002/rob.21699.
Bresilla, K., Perulli, G.D., Boini, A., Morandi, B., Corelli Grappadelli, L., Manfrini, L., 2019. Single-shot convolution neural networks for real-time fruit detection within the tree. Front. Plant Sci. 10, 611. https://doi.org/10.3389/fpls.2019.00611.
Cheng, H., Damerow, L., Sun, Y., Blanke, M., 2017. Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks. Journal of Imaging 3 (1), 6. https://doi.org/10.3390/jimaging3010006.
Christensen, J.V., 1979. Effects of density, rectangularity and row orientation on apple trees, measured in a multivariated experimental design. Scientia Horticulturae 10 (2), 155–165. https://doi.org/10.1016/0304-4238(79)90121-3.
Cripps, J.E., 2016. Apple tree named 'ANABP 01.' USPP27420P3.
Devyatov, A.S., Gorny, A.V., 1978. Effect of espalier orientation on light status andcropping of apple trees. Fruit Sci. Rep. 5, 1–8.
Dias, P.A., Tabb, A., Medeiros, H., 2018a. Multispecies Fruit Flower Detection Using a Refined Semantic Segmentation Network. IEEE Robot. Autom. Lett. 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. Comput. Ind. 99, 17–28. https://doi.org/10.1016/j.compind.2018.03.010.
Gómez-del-Campo, M., Centeno, A., Connor, D.J., 2009. Yield determination in olive hedgerow orchards. I. Yield and profiles of yield components in north–south and

east–west oriented hedgerows. *Crop and Pasture*. Science 60 (5), 434–442. https://doi.org/10.1071/cp08252.

Goodwin, I., Whitfield, D.M., Connor, D.J., 2006. Effects of tree size on water use of peach (*Prunus persica* L. Batsch). Irrig. Sci. 24 (2), 59–68. https://doi.org/10.1007/s00271-005-0010-z.

Hunter, J.J., Volschenk, C.G., Booyse, M., 2017. Vineyard row orientation and grape ripeness level effects on vegetative and reproductive growth characteristics of *Vitis vinifera* L. cv. Shiraz/101-14 Mgt. *European Journal of* Agronomy 84, 47–57. https://doi.org/10.1016/j.eja.2016.12.004.

Kang, H., Chen, C., 2020. Fast implementation of real-time fruit detection in apple orchards using deep learning. Comput. Electron. Agric. 168, 105108. https://doi.org/10.1016/j.compag.2019.105108.

Kuznetsova, A., Maleva, T., Soloviev, V., 2020. Using YOLOv3 Algorithm with Pre- and Post-Processing for Apple Detection in Fruit-Harvesting Robot. Agronomy 10, 1016. https://doi.org/10.3390/agronomy10071016.

Lin, L.-I.-K., 1989. A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics 45, 255–268. https://doi.org/10.2307/2532051.

McClymont, L., Goodwin, I., Mazza, M., Baker, N., Lanyon, D.M., Zerihun, A., Chandra, S., Downey, M.O., 2012. Effect of site-specific irrigation management on grapevine yield and fruit quality attributes. Irrigation Science 30 (6), 461–470. https://doi.org/10.1007/s00271-012-0376-7.

McClymont, L., Goodwin, I., O'Connell, M.G., Whitfield, D.M., 2011. Variation in within-block orchard canopy cover – implications for crop water requirement and irrigation management. Acta Hortic. 889, 233-239, 10.17660/ActaHortic.2011.889.27.

Middleton, S., McWaters, A., James, P., Jotic, P., Sutton, J. and Campbell, J., 2002. The productivity and performance of apple orchard systems in Australia. Compact Fruit Tree, 35(2), pp.43-47.

Örn, S., 2016. Estimating Light Interception of Orchard Trees Using LiDAR and Solar Models. Master Degree Thesis, Linköping University.

Palmer, J.W., 1989. The effects of row orientation, tree height, time of year and latitude on light interception and distribution in model apple hedgerow canopies. Journal of Horticultural Science 64 (2), 137–145. https://doi.org/10.1080/14620316.1989.11515937.

Peters G. Userfriendlyscience: Quantitative Analysis Made Accessible. 2018. [accessed on 1 May 2021]. R Package Version 0.7.2. Available online: https://userfriendlyscience.com.

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2018. [accessed on 1 May 2021]. Available online: https://www.R-project.org/.

Robinson, T. L. and Lakso, A., 1989. Light interception, yield and fruit quality of 'Empire' and 'Delicious' apple trees grown in four orchard systems. Acta Hortic. 243, 175-184. 10.17660/ActaHortic.1989.243.21.

Signorell, A. DescTools: Tools for Descriptive Statistics and Exploratory Data Analysis. 2021. [accessed on 1 May 2021]. R Package Version 0.99.41. Available online: https://github.com/AndriSignorell/DescTools/.

Trentacoste, E.R., Connor, D.J., Gómez-del-Campo, M., 2015. Row orientation: applications to productivity and design of hedgerows in horticultural and olive orchards. Scientia Horticulturae 187, 15–29. https://doi.org/10.1016/j.scienta.2015.02.032.

Underwood, J.P., Jagbrant, G., Nieto, J.I., Sukkarieh, S., 2015. Lidar-Based Tree Recognition and Platform Localization in Orchards. J. F. Robot. 32, 1056–1074. https://doi.org/10.1002/rob.21607.

Underwood, J.P., Hung, C., Whelan, B., Sukkarieh, S., 2016. Mapping almond orchard canopy volume, flowers, fruit and yield using lidar and vision sensors. Comput. Electron. Agric. 130, 83–96. https://doi.org/10.1016/j.compag.2016.09.014.

Univer, T., Univer, N. and Tiirmaa, K., 2017, May. Results of a study of the vegetative growth, yield and fruit quality of different grafting combinations of the apple cultivar 'Krista'. In Proceedings of the Latvian Academy of Sciences (Vol. 71, No. 3, p. 137). De Gruyter Poland. 10.1515/prolas-2017-0024.

Wagenmakers, P.S., Callesen, O., 1995. Light distribution in apple orchard systems in relation to production and fruit quality. Journal of Horticultural Science 70 (6), 935–948. https://doi.org/10.1080/14620316.1995.11515369.

Wang, Z., Underwood, J., Walsh, K.B., 2018. Machine vision assessment of mango orchard flowering. Comput. Electron. Agric. 151, 501–511. https://doi.org/10.1016/j.compag.2018.06.040.

Westling, F., Underwood, J., Örn, S., 2018. Light interception modelling using unstructured LiDAR data in avocado orchards. Comput. Electron. Agric. 153, 177–187. https://doi.org/10.1016/j.compag.2018.08.020.

Westling, Fred, Mahmud, Kare, Underwood, James, Bally, Ian, 2020. Replacing traditional light measurement with LiDAR based methods in orchards. Comput. Electronics Agriculture 179, 105798. https://doi.org/10.1016/j.compag.2020.105798.

Westling, Fred, Underwood, James, Bryson, Mitch, 2021. A procedure for automated tree pruning suggestion using LiDAR scans of fruit trees. Comput. Electr. Agriculture 187, 106274. https://doi.org/10.1016/j.compag.2021.106274.

Wünsche, J.N. and Lakso, A.N., 2000. The relationship between leaf area and light interception by spur and extension shoot leaves and apple orchard productivity. HortScience, 35(7), pp.1202-1206. 10.21273/hortsci.35.7.1202.

Wünsche, J. N., Lakso, A. N., Robinson, T. L., Lenz, F., and Denning, S. S., 1996. The bases of productivity in apple production systems: the role of light interception by different shoot types. Journal of the American Society for Horticultural Science, 121 (5), 886-893. 10.21273/JASHS.121.5.886.