CrossMark

# Estimation of fruit load in mango orchards: tree sampling considerations and use of machine vision and satellite imagery

N. T. Anderson[1] · J. P. Underwood[2] · M. M. Rahman[3] · A. Robson[3] · K. B. Walsh[1]

## Abstract

In current best commercial practice, pre-harvest fruit load predictions of mango orchards are provided based on a manual count of fruit number on up to 5% of trees within each block. However, the variability in fruit number per tree (coefficient of variation, CV, from 27 to 93% across ten orchards) was demonstrated to be such that the best case commercial sampling practice was inadequate for reliable estimation (to an error of 54–82 fruit/tree, and percentage error, PE, of 10% at a probability of 0.95). These results highlight the need for alternative methods for estimation of orchard fruit load. Pre-harvest fruit load was estimated for a case study orchard of 469 trees using (i) count of a sample of trees, (ii) in-field machine vision and (iii) correlation to a tree spectral index estimated using high resolution satellite imagery. A count of 5% of trees (23) in the trial orchard resulted in a PE of 31% (error of 37 fruit/tree), with a count of 157 trees required to achieve a PE of 10% (error of 12 fruit/tree). Sampling effort to achieve a PE of 10% was decreased by only 10% by sampling from aspatial k-means tree classifications based on machine vision derived fruit counts of all trees. Clustering based on tree attributes of canopy volume and trunk circumference was not helpful in decreasing sampling effort as these attributes were poorly correlated to fruit load ($R^2 = 0.21$ and 0.17, respectively). In-field multi-view machine vision-based estimation of fruit load per tree achieved a $R^2 = 0.97$ and a RMSE = 14.8 fruit/tree against harvest fruit count per tree for a set of 18 trees (average = 88; SD = 82 fruit/tree), using a faster region convolutional neural network trained the previous season. The relationship between WorldView-3 (WV3) satellite spectral reflectance characteristics of sampled trees and fruit number was characterised by a $R^2 = 0.66$ and a RMSE = 56.1 fruit/tree. For this orchard, for which the actual fruit harvest was 56,720 fruit, the estimate based on a manual count of 5% of trees was 47,955 fruit, while estimates based on 20 iterations of stratified sampling (of 5% of trees in each cycle) had variation (SD) of 9597. The machine vision method resulted in an estimate of 53,520 (SD = 1960) fruit and the remote sensing method, 51,944 (SD = 26,300) fruit for the orchard.

**Keywords** Fruiticulture · Getis–Ord · Precision horticulture · Remote sensing · Tree crop · Yield estimation · Machine vision

✉ N. T. Anderson
  nicholas.anderson@cqumail.com

Extended author information available on the last page of the article

🖄 Springer

## Introduction

Accurate pre-harvest estimations of fruit load per tree can support more informed decisions regarding harvesting logistics (e.g. labour, equipment and packing material requirements) as well as handling, storage and forward selling. This requirement is particularly strong in mango (*Mangifera indica L.*) production, given the narrow window of time between harvest maturity and ripening on tree. Managers of smaller mango farms rely on a qualitative assessment of fruit load (e.g. the block appears to have 20% more or less fruit than last year). Larger farms utilize a quantitative estimation based on a visual count of a sub-sample of trees from each tree block, in best practice involving a count of fruit on every 20th tree (i.e. 5% of trees; 50 trees for a block of 1000 trees) (pers. comm. M. Matzner, farm manager). In some cases, an average of three counts per block is made, changing the initial count tree between counts (pers. comm. M. Robertson, farm manager). This procedure will be beneficial if there is a high variation between trees, however there is no published information on mango yield variability in the context of yield forecasts.

More accurate crop load estimations with lower sampling effort could potentially be achieved if an orchard was segregated into clusters of trees with similar crop load. This could be undertaken if a correlation existed with an easily measured field attribute and fruit load per tree. For example, mango trees produce inflorescences at branch terminals, and thus trees with more branch terminals have a greater potential fruit load. Therefore the potential fruit load per tree could be proportional to canopy volume, which in turn could be proportional to trunk circumference or foliage health as indexed by a spectral index.

An attempt was made to optimise human sampling effort by Taylor et al. (2007) who mapped spatial variability in kiwifruit weight and dry matter content over three harvest seasons across 11 orchards to determine a minimum sampling grid for a desired confidence level. More recently, Peeters et al. (2015) reported a technique for combining spatial and aspatial information to classify variation within orchards. The attributes of soil electrical conductivity, trunk circumference and yield were used in development of a classified map using the aspatial method of k-means clustering to classify trees to separate clusters, while the Getis–Ord ($G_i^*$) univariate geospatial analysis method (Getis and Ord 1992) was used to include spatial information by including reference to neighbouring tree data values. Individual tree spatial significance scores from the $G_i^*$ analysis were used as input variables for the aspatial k-means clustering analysis, resulting in a classification with spatial structure.

There is also potential for yield estimation of tree fruit crops using machine vision or remote sensing. If low cost and easily implemented (i.e., standardised and automated), such technologies could supplant manual counts of tree fruit load. At a lower level of adoption, these technologies could be used for a qualitative classification of trees by level of crop load (e.g. high, medium, low), allowing for a reduced manual sampling effort.

Machine vision systems have been previously examined for their ability to accurately estimate mango fruit load on a tree. Payne et al. (2013) reported the use of a night imaging, dual-view approach (two images of each tree, from the two inter-rows), with a coefficient of determination ($R^2$) of 0.74 and bias corrected root mean square error (RMSE) of 13.3 fruit/tree (RMSE recalculated from presented data) for the regression of observed fruit for the estimation of fruit number per tree compared to actual count. The dual-view count underestimated actual fruit number due to canopy occlusion, with only 59 and 84% of total fruit visible in dual-view images in two separate populations. Stein et al. (2016) used a faster regional convolutional neural network (R-CNN) detector with both a dual-view and a multi-view per tree approach that employed trajectory data, tracking individual fruits

between frames to avoid repetitive counts. A LiDAR image mask was used to identify individual canopies to associate detected fruits with individual trees. The multi-view count versus a harvest count for 16 validation trees was both precise ($R^2 = 0.90$) and accurate (slope $= 1.01$, with double counts balanced by hidden fruit) while the dual-view approach was more precise ($R^2 = 0.94$) but less accurate (slope $= 0.54$). Denser canopies could prevent a view of all fruit on the canopy, such that a correction factor for hidden fruit would be required for the machine vision count. Thus implementation of a machine vision-based fruit load estimator may still require manual sampling effort.

In the only reported use of satellite imagery in the context of mango production, Yadav et al. (2002) reported on the use of the Indian Remote Sensing (IRS) satellite for estimation of the production area of mango. The authors also attempted to develop a multiple linear regression model for yield prediction based on field collected canopy attributes (tree height, scion trunk diameter, canopy width) on 20 treatments of 10 trees each over 9 seasons (n $= 1800$). Weak correlations were reported between individual tree attributes and yield ($R^2 < 0.36$), and between all attributes and yield (multiple linear regression, $R^2 = 0.53$).

The use of 18 vegetation indices (VI) estimated from WV3 satellite multispectral imagery as predictors of yield of avocado and macadamia blocks was reported by Robson et al. (2017). Reasonable linear correlations between a canopy VI and fruit yield (kg/tree) were achieved for two macadamia and three avocado blocks ($R^2 = 0.86$, 0.69, 0.81, 0.68, 0.72; respectively; data of 18 trees per block). Inconsistencies in the correlation were explained (e.g. macadamia nut losses due to a hail event). As well as providing information on variability between trees within an orchard, the remote sensing approach allows for a yield forecast for entire growing districts. However, the specific VI employed and the relationship slope between a given VI and fruit load varied between blocks. This variation may be attributed to seasonal, locational and management differences (resulting in variation in flowering extent and fruit set per unit of canopy volume). Thus, to apply this methodology, field-based measurements will be required to calibrate the relationship between VI and fruit load.

In the current study, the reliability of manual crop load estimation was assessed, with an attempt to use the approach of Peeters et al. (2015) to reduce sampling effort based on use of attributes to classify trees to groups with reduced variation in fruit load. Further, the accuracy of crop load estimates using manual count, machine vision and remote sensing methods were compared.

## Materials and methods

### Field material and harvest

Field work was undertaken in ten orchards in the 2016/17 season. The main study site (Orchard 1) was a commercial mango (cultivar 'Calypso™') block near Bundaberg, Australia (around 24.8670°S, 152.3510°E). This site was also utilised in the study of Stein et al. (2016) in a previous year. The trees were planted on a $4 \times 9$ m grid, and were 12 years old. Of the 491 trees originally planted in the orchard, 22 were cut back to major branches and top-worked with grafts in the previous season, while another ten trees bore no fruit.

In Orchard 1, fruit ($> 225$ g) was harvested and counted from 18 'calibration' trees selected on the basis of canopy VI (six each of high, medium and low values; assessed

using WV3 satellite imagery, see description in next section), with harvest on the day before commercial harvest. The entire block was commercially harvested on the 16th of January 2017, with fruit sorted using a Compac (Auckland, NZ) fruit grader, with total orchard yield (# of fruit > 225 g) recorded. Fruit < 225 g was typically very small (approximately 50 g), being fruit without seed, derived from non-fertilised flowers.

In the other nine orchards (2–10), 18 trees were selected and fruit load was counted at harvest, as described above. These orchards were located in the Northern Territory (around 12.5753°S, 131.1022°E) and Queensland (around 25.2370°S, 152.2685°E), Australia, and were of Calypso™, Honey Gold™, Kensington Pride and R2E2 cultivars.

## Measurements

A visual count of fruit on a tree requires the operator to work systematically around the tree, making counts by zones (generally branches). Manual visual counts of fruit on tree were made of 191 trees in Orchard 1, including the 18 calibration trees. A repeat visual fruit count of fruit on 18 trees was made by two trained operators to estimate measurement error. Trunk circumference (0.1 m above the graft union) of all Orchard 1 trees was measured using a measuring tape.

WV3 satellite imagery with 1.2 m spatial resolution of Orchard 1 was obtained on September 23rd, 2016, and the method of Robson et al. (2016, 2017) employed. Briefly, pixels specific to tree canopies were segmented using a 2D scatter plot (Red versus NIR1) in ENVI version 5.4 (Exelis Visual Information Solutions, Boulder, Colorado, USA). The Normalised Difference Vegetation Index (NDVI = (NIR1 − R)/(NIR1 + R), where NIR1 is near infrared band 1, 772–890 nm, and R is red band, 632–692 nm) was calculated and used in an unsupervised classification, using Iso Cluster and Maximum Likelihood Classification tools (ArcGIS 10.2), to high, medium and low NDVI categories. Six trees were randomly selected from each class for calibration activities. Pixels associated with each of the 'calibration' trees were segmented by applying a 1.5 m radius buffer area around each central point of the tree using ArcGIS 10.2 (Environmental Systems Research Institute, Redlands, CA, USA). Eighteen structural and pigment based VIs specific to crop biomass and yield parameters were derived for each tree from the eight band spectral reflectance data of WV3 imagery (Robson et al. 2017). Tree fruit load was regressed against the 18 VIs for the 'calibration' trees. The VI returning the highest regression coefficient of determination for the orchard was adopted, with the regression equation applied to the average VI of canopy associated pixels of the whole orchard. The estimated average yield per tree for the orchard was multiplied by tree number to provide an estimate of orchard fruit yield.

The techniques involved in RGB-LiDAR-based estimates of canopy volume and multi-view machine vision estimates of fruit load per tree were described by Stein et al. (2016) based on field work in Orchard 1 in the previous season. RGB and LiDAR imagery were collected of all trees in the orchard within December, 2016. Briefly, imagery was obtained from a ground-based platform of all trees in the orchard. A faster region convolutional neural network (R-CNN) algorithm was trained on > 1500 images of individual fruit, and the method was validated using the manual (harvested) counts of fruit from the 18 trees mentioned earlier. The technique was then run on full-sized images for all trees in the orchard, both for a single image per side of the tree ('dual-view') and for tracked fruit in 25 images per side of each tree ('multi-view'). Fruit detections within each image were associated with a specific tree by projecting segmented LiDAR data of each tree to the corresponding image frames. For the current study, the Stein et al. (2016) method was updated to clip the

LiDAR masks at a vertical plane, intersecting the centroid of the canopy in the direction of the row, to avoid double counting fruit that were seen from both sides of the tree. The faster R-CNN was used without further training from the previous season. The method was validated using the manual (harvested) counts of fruit from the 18 trees mentioned earlier, and using the manual on-tree fruit counts of 191 trees. The slope of the relationship between machine vision estimates and actual count of the 18 calibration trees was used to adjust estimates. The slope for the multi-view relationship was effectively unity, at 1.0043. The summation of machine vision tree fruit load estimates provided an orchard yield estimate. Individual tree load multi-view estimates were used in consideration of the reliability of sub-sampling estimates.

## Statistical analysis

For convenience, the terms used in the equations of this study are listed here:

$n$ is the number of sample trees; $N$ is the total number of trees in the orchard, $n_{adj}$ is adjusted sampling requirement (Eq. 1); $t$ is the t statistic, CV is coefficient of variation, $e$ is measurement error, PE is percentage error (Eq. 2); $c$ is the machine vision count of fruit per tree, $\hat{c}$ is the corrected machine vision count of fruit per tree, $m$ is the slope of the actual (harvest) count, $\bar{c}$, to machine vision count, with associated uncertainty of $\sigma_m$ (Eqs. 3, 4); VI is satellite image derived spectra vegetation index, $z$ is the slope of $\bar{c}$ to VI and $z$ is the intercept of this relationship (Eq. 5); $X_i$ is a range normalised tree variable, $x_i$ is the value for the $i$th tree for a given value (Eq. 6); $X_j$ is the variable value of tree $j$, $w_{b,j}$ is the spatial weight between attributes $b$ and $j$ (Eq. 7); $k$ is the number of clusters used, $s_i$ represent a single cluster within $k$, $X_N$ is the variable per $N$, and $\mu_i$ is the mean of $X_N$ within $s_i$ (Eq. 8)

The minimum number of samples (tree counts) required ($n$) for a reliable estimate of population mean (of fruit per tree) was calculated from the standard deviation (SD), in the context of the desired probability level (using a $t$ value; Students $t$ table: for P of 0.95 and $n > 30$, $t = 1.96$) and the acceptable measurement error ($e$) (Eq. 1). Alternatively, CV (SD/mean * 100%) and PE ($e$/mean *100%) can be used (Eq. 1).

$$n = \frac{(t \times SD)^2}{e^2} = \frac{(t \times CV)^2}{PE^2}. \tag{1}$$

For finite populations with high SD, required sampling $n$ estimated from Eq. 1 can (unreasonably) exceed population size. An adjusted sampling requirement, $n_{adj}$ (Thrusfield 1995) can be calculated given the total population ($N$, trees per orchard) as:

$$n_{adj} = \frac{n \times N}{N + n}. \tag{2}$$

The impact of varying the 'start' tree in the commercial yield estimation practice of assessing every 20th tree was assessed using the multi-view machine vision-based tree fruit load data. Yield estimates were made based on the start tree varying from tree 1 to 20 (where tree 1 was the edge tree of the orchard north-west corner).

The machine vision fruit count estimates were corrected based on a Bayesian linear regression of machine vision estimates ($c$) against harvested fruit counts ($\bar{c}$) for the calibration trees. Observation and model prior noise ($\sigma_o$ and $\sigma_m$) were optimised to maximise the marginal likelihood (Rasmussen and Williams 2006), to estimate a line through the origin with a Gaussian uncertainty distribution on the gradient. Observation noise $\sigma_o$ is equivalent to the RMSE of the calibration data. The prior distribution on the slope ($m$) is a Gaussian centred at 0 with

standard deviation $\sigma_m$. Machine vision counts per tree $c$ were converted to count estimates $\hat{c}$ per tree (Eq. 3):

$$\hat{c} = \frac{c}{(m \pm \sigma_m)}, \tag{3}$$

where the uncertainty in calibration gradient $m$ is described by the Gaussian $\mathrm{N}(m, \sigma_m^2)$ with mean gradient $m$ and standard deviation $\sigma_m$. Block total estimates for $N$ trees were calculated by:

$$total = \frac{\sum c}{m} \pm \left( \sqrt{N}\sigma_o + \left( \frac{\sum c}{m + \sigma_m} - \frac{\sum c}{m - \sigma_m} \right) \bigg/ 2 \right). \tag{4}$$

In this expression, the error term has two components: the first $(\sqrt{N}\sigma_o)$ represents repeated samples of the observation noise for every tree, while the second $\left( \left( \frac{\sum c}{m+\sigma_m} - \frac{\sum c}{m-\sigma_m} \right) \bigg/ 2 \right)$ is derived from the upper and lower estimates of the calibration slope (plus and minus one standard deviation).

For WV-3 data, a linear regression of VI to harvested fruit number per tree ($\bar{c}$) was rearranged (Eq. 5, where $z$ is the intercept and $m_1$ is the slope of the regression line) for use in conversion of the average VI of orchard canopy pixels to an estimate of the average tree fruit number for the orchard. RMSE of the predicted fruit count/tree to actual count for the calibration set was used in estimation of error on the estimate of orchard fruit count as RMSE/($c$) multiplied by the number of trees in the orchard (n = 469).

$$\bar{c} = \frac{(VI - z)}{m1}. \tag{5}$$

## Spatial clustering

The spatial clustering procedures followed protocol of ArcGIS (Esri 2018). Each tree level attribute was pre-processed by standardization (0–1) (Eq. 6) prior to further analysis.

$$X_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}, \tag{6}$$

where $x = (x_i,\ldots,x_N)$ and $Xi$ is the standardized value.

Clustering techniques were used independently for each tree attributes (i.e., machine vision derived fruit numbers, canopy volume and trunk circumference) and in a multivariate assessment involving all three attributes. To include spatial information, the methodology of Peeters et al. (2015) was followed, in which z-scores from the univariate spatial Getis–Ord were used as inputs for the multivariate aspatial k-means clustering. This procedure allows spatial data (z-scores) to be used in an aspatial clustering method.

The standardized $G_i^*$ (Eq. 7), developed by Ord and Getis (1995), determines if a data point is statistically different to neighbouring objects (plotted at P < 0.01, 0.05, 0.10 assuming a normal distribution). In this expression, $w_{b,j}$ is the spatial weight between features $b$ and $j$ within a binary (0–1) symmetric spatial weight matrix, with a value of 1 assigned to neighbouring points (adjacent trees) and a value of 0 assigned to all other points and including the point $b$. $N$ is the total number of data points (i.e. trees) and $X_j$ is the value for the data point $j$.

$$G_i^* = \frac{\sum_{j=1}^{N} w_{b,j} X_j - \bar{X} \sum_{j=1}^{N} w_{b,j}}{s\sqrt{\dfrac{\left[N \sum_{j=1}^{N} w_{b,j}^2 - \left(\sum_{j=1}^{N} w_{b,j}\right)^2\right]}{N-1}}}, \tag{7}$$

where:

$$\bar{X} = \frac{\sum_{j=1}^{N} X_j}{N}; \text{ and } S = \sqrt{\frac{\sum_{j=1}^{N} X_j^2}{N} - (\bar{X})^2}.$$

The k-means algorithm involves random placement of centroids ($s$) for a given number of clusters ($k$) (Eq. 8).

$$K = \arg\min \sum_{i=1}^{k} \sum_{X_N \in s_i} \left\| X_N - \mu_i \right\|^2, \tag{8}$$

where arg*min* is the minimum Euclidean distance between tree attribute $(X_N)$ and the mean value $(\mu_i)$ of all points within the given centroid $(s_{i_j})$. After trees are assigned to clusters, the mean of all point values within a given cluster changes, thus the procedure was repeated until there were no changes in tree assignments to clusters.

The optimal number of clusters was determined using RStudio (Boston, USA), based on the intersection in a plot of the within-cluster sums of squares (WCSS) and the between-clusters sums of squares (BCSS) for a range in the number of clusters. For each tree variable, $G_i^*$ analysis was performed in ArcGIS to derive spatial significance values (z-score) for each data point. The z-scores for each variable were then used as the input values for the k-means clustering ('Grouping Analysis'). ArcGIS was used to plot raw, k-means clustered and Getis–Ord $G_i^*$ spatially clustered values.

**Table 1** Average fruit number and standard deviation for 2016/17 harvests of 18 trees in each of 10 mango orchards (1–10) and for the multi-view machine vision estimate of all trees in Orchard 1 (1*)

| Orchard | Tree N (#/orchard) | Average (fruit/tree) | SD (fruit/tree) | CV (%) | Sampling n (trees) | Adjusted n (trees) |
|---|---|---|---|---|---|---|
| (1*) | 469 | 117 | 92 | 78 | 235 | 157 |
| 1 | 469 | 88 | 82 | 93 | 334 | 195 |
| 2 | 486 | 259 | 102 | 39 | 60 | 53 |
| 3 | 1017 | 240 | 160 | 67 | 171 | 146 |
| 4 | 1100 | 80 | 34 | 43 | 69 | 65 |
| 5 | 224 | 59 | 36 | 61 | 143 | 87 |
| 6 | 1205 | 97 | 65 | 67 | 173 | 151 |
| 7 | 1091 | 201 | 55 | 27 | 29 | 28 |
| 8 | 1818 | 106 | 51 | 48 | 89 | 85 |
| 9 | 1176 | 77 | 61 | 79 | 241 | 200 |
| 10 | 1117 | 85 | 40 | 47 | 85 | 79 |

A PE of 10% and P of 0.95 was used to determine a minimum number of samples (n) required for estimation of the mean for each orchard (for P = 0.95)

## Results and discussion

### Minimum sample number for estimate of orchard yield

The variation in fruit number per tree was high in all sampled orchards, with SD ranging from 34 to 160 fruit/tree (CV of 27 to 93%) (Table 1). These results are consistent with those of Payne et al. (2013) who reported SD of 37 and 50 fruit/tree for two orchards (CV = 44 and 56%; average = 84 and 93 fruit/tree; respectively) and generally higher than that reported for a grapefruit orchard (CV = 42%) (Peeters et al. 2015). The adjusted number of samples required for an estimate of the mean tree fruit load (using Eq. 2; for P = 0.95, PE = 10%) ranged from 28 to 200 trees across the 10 orchards.

For the main study orchard (Orchard 1; average = 117 fruit/tree; SD = 91.6 fruit/tree on multi-view count), the number (adjusted n) of sample trees to achieve an estimate of the mean tree crop load for the block (for P = 0.95) varied from 314 for a 5% measurement error to 23 trees if a 31% error was accepted (5% of total tree number of this block, n = 469) (Fig. 1). To achieve a 'reasonable' PE of 10%, a count of 157 (235 non-adjusted) trees was required (Table 1). Random selection (n = 20) of (dual view estimated) values of 157 trees resulted in estimates of between 106 and 129 fruit/tree, with average = 107 and SD = 6.1 fruit/tree (i.e. CV of 5.6%).

Stratified sampling is the method of sampling every $n$th tree, typically resulting in diagonal sampling 'lines' across a rectangular orchard with rows of equally spaced trees. To demonstrate the limitation of a 5% of total sampling method, the start tree for an 'every 20th tree' stratified sample was varied. As expected, the SD of these estimates (20.5 fruit/tree) was decreased relative to the SD (92 fruit/tree) of individual tree fruit counts across the orchard. However, the variation in average tree yield and thus orchard yield based on the stratified sampling estimates was unacceptable in terms of using a 5% sampling strategy for a commercially useful estimate, with tree yield estimates varying from 79 to 145 fruit/tree (CV = 18%, Table 2).
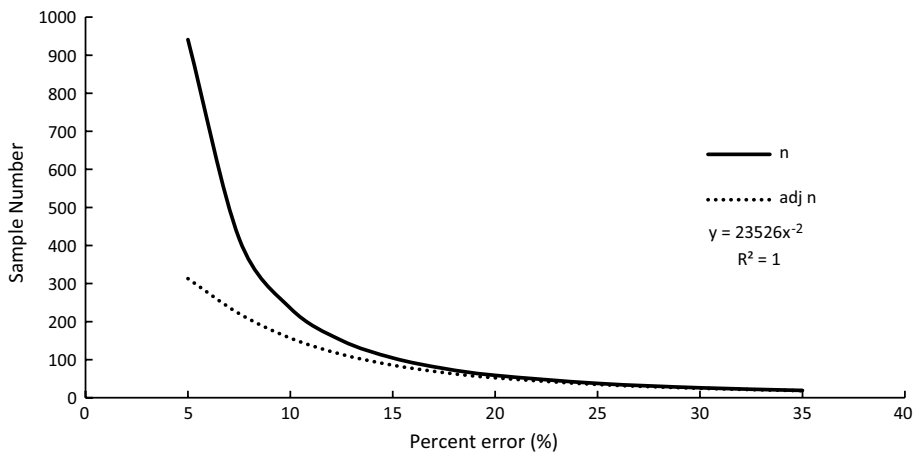


**Fig. 1** Number of samples (n) required to estimate population mean at P = 0.95 in context of accepted PE (%) for a population, given a CV of 78.7%. Results shown for raw n and adjusted n

| Initial tree | Average fruit/tree | Estimated orchard fruit # |
|---|---|---|
| 1 | 122.8 | 57608 |
| 2 | 127.8 | 59953 |
| 3 | 112.2 | 52645 |
| 4 | 145.3 | 68161 |
| 5 | 118.2 | 55478 |
| 6 | 98.3 | 46118 |
| 7 | 123.1 | 57765 |
| 8 | 92.6 | 43441 |
| 9 | 130.8 | 61360 |
| 10 | 129.9 | 60908 |
| 11 | 144.9 | 67943 |
| 12 | 79.0 | 37071 |
| 13 | 117.3 | 55036 |
| 14 | 90.2 | 42291 |
| 15 | 104.3 | 48898 |
| 16 | 108.7 | 50978 |
| 17 | 82.5 | 38702 |
| 18 | 82.8 | 38845 |
| 19 | 119.3 | 55953 |
| 20 | 136.4 | 63987 |
| Mean | 113.3 | 53158 |
| SD | 20.5 | 9597 |
| CV (%) | 18.1 | 18.1 |

**Table 2** Average fruit counts per tree and estimated orchard yield based on data of every 20th tree, with variation in start tree used in the count

## Tree crop load estimates

For Orchard 1, repeat manual estimates of fruit load on the calibration trees (i.e. pre harvest) achieved an $R^2 = 0.998$, RMSE of 2.61 fruit/tree and a slope $= 1.01$. This RMSE value represents an estimate of the measurement uncertainty of a (trained) human visual estimate working under ideal conditions. The canopies of the trees in Orchard 1 were sufficiently open that manual count could be accurate and repeatable. The accuracy of a human count is expected to decrease with count of large tree numbers (operator attention span limitation), time pressure (commercial reality), higher fruit loads per tree and larger canopies (data not shown).

Of the 18 spectral indices calculated from the satellite imagery, the best relationship was obtained with the N2RENDVI index ((NIR2 − R)/(NIR2 + R)). A $R^2$ of 0.66 was obtained for the linear correlation of VI to fruit load estimated by harvest for 15 'calibration' trees of Orchard 1 (the three trees in the calibration set that had no fruit were not utilised in the regression; Fig. 2). An RMSE of 56.1 fruit/tree was calculated for the fruit count estimated using this regression relative to harvest count.

For machine vision estimates of fruit load per tree, Stein et al. (2016) reported that for the 'calibration' trees in the previous season (n = 16, average = 114, SD = 79 fruit/tree), the linear regression of manual, dual and multi-view machine vision counts on harvest counts
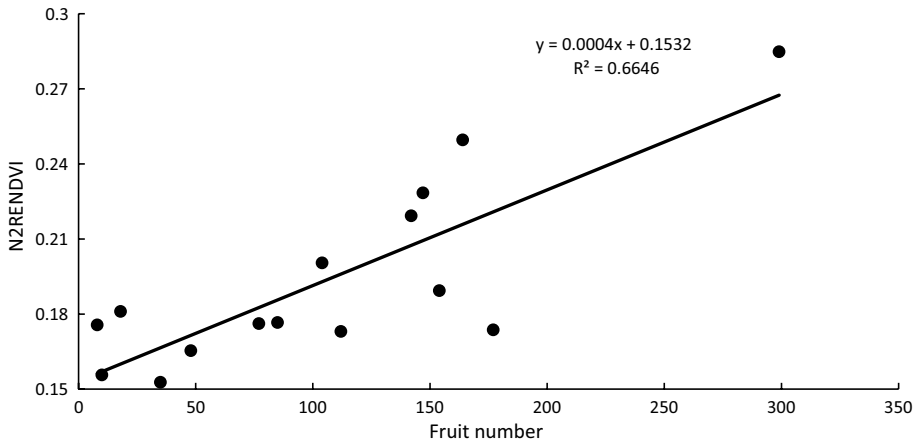
**Fig. 2** Plot of fruit number and NIRENDVI for calibration trees of Orchard 1

was described by a $R^2$ of 0.99, 0.94 and 0.90, and associated slopes of 1.03, 0.54, and 1.01, respectively. The machine vision algorithms developed by Stein et al. (2016) in the previous season were used in the current study, with addition of a 'half-mask' clip such that only fruits in the half of the canopy facing the imaging platform were counted. The linear regression of LiDAR masked dual and multi-view machine counts, respectively, on harvest counts of the calibration trees (n = 18, average = 88; SD = 82 fruit/tree) was described by a $R^2$ of 0.90 and 0.97, RMSE of 25 and 15 fruit/tree and slope of 0.50 and 1.0043. In comparison, the result for dual-view without use of a mask (i.e. whole image assessed) was described by a $R^2$ of 0.81, RMSE of 30 fruit/tree and slope of 0.74. The reduction in $R^2$ is due to misallocation of fruit seen in neighbouring trees, which affects calibration but not total block estimates. The high precision and accuracy of the multi-view counts of Orchard 1 is consistent with the observation that most fruit could be seen from some angle from outside the canopy given the relatively open tree canopies. The multi-view counts were therefore used in consideration of required sampling effort and potential clustering procedures to reduce manual sampling effort.

Human visual assessment of fruit number on tree was taken within days of orchard imaging for machine vision assessment, while harvest occurred a month later. In this season, there was effectively no fruit drop in this period. The linear regression of multi-view machine vision counts on human visual assessment of tree load for 191 trees (average = 95.9; SD = 61.8 fruits/tree) was described by a $R^2 = 0.85$, RMSE = 35.8 fruit/tree and slope = 1.2. The poorer values from human estimates of the 191 tree set than for the harvest data of the 18 'calibration' trees is attributed to error in the human in-field count, representing operator fatigue for assessment of larger sample sizes.

## Orchard level variation

In Orchard 1, trees in the north-western corner of the block tended to have higher values of fruit load, trunk diameter and canopy volume (Fig. 3), while trees in the mid-east of the block had lower values. However, trunk circumference was weakly linearly correlated with canopy volume ($R^2 = 0.57$), presumably due to variation in pruning operations between trees. Further, fruit load per tree was poorly linearly correlated to either (LiDAR estimated)
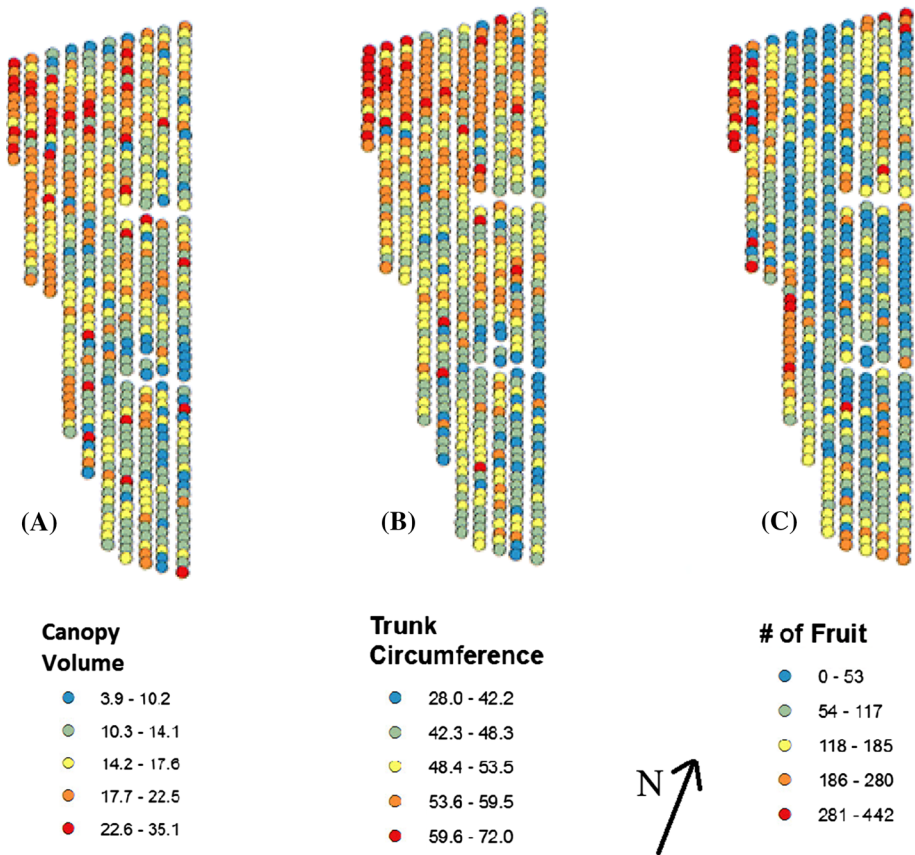
**Fig. 3** Orchard map with trees (n=469) colour coded for canopy volume (**a**; m³), trunk circumference (**b**; cm) and number of fruits per tree (**c**). Each circle represents a tree (Color figure online)

canopy volume or (manually estimated) trunk circumference (R²=0.21 and 0.17, respectively). While canopy volume could relate to the number of terminals (branches) per tree and thus to the potential fruit load per tree, the poor relationship between volume and yield must reflect variation in the percentage of terminals that flower or in % fruit drop within this orchard and season.

Peeters et al. (2015) recommended the variables of trunk circumference (as a surrogate for canopy volume) and tree fruit load for orchard classification into management units, but they did not report on the correlation between trunk circumference and yield. Given the poor correlation between canopy volume or trunk circumference and yield in the current exercise, orchard classification based on these variables should have little relevance to yield, but may have value for other management purposes (e.g., pruning effort or fertilisation).
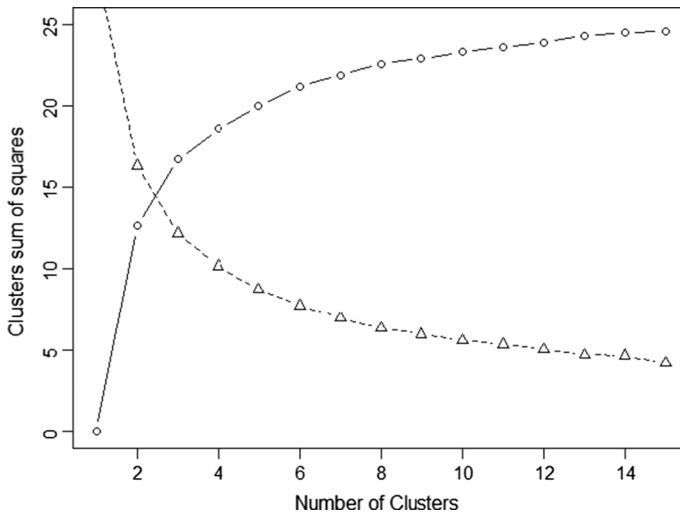
**Fig. 4** Plot of within class sum of squares (triangles) and between class sum of squares (circles) against number of clusters for a k-means classification of 469 trees on canopy volume, trunk circumference and fruit number per tree

## Orchard classification

Orchard classification based on prior knowledge of tree fruit load was undertaken, to assess the decreased sampling effort possible through measurement within zones of lower variation in fruit load. Classification of trees by the indices of canopy volume and trunk circumference was also undertaken, to replicate the work of Peeters et al. (2015).

For all variables, the k-means optimization plot of within cluster and between cluster sums of squares intersected between two and three clusters (e.g., Fig. 4). Three clusters represent a practical number of groups in terms of interpretation (high, medium, low), and this number was adopted for all classifications.

The k-means classification was consistent with visual assessment of the distribution of fruit/load across the orchard, i.e. higher fruit loads in the north-western corner of the block (Fig. 5). This distribution was also evident in the $G_i^*$ hot-spot analysis (Fig. 5).

A multivariate classification was attempted using the three variables (trunk diameter, canopy volume and crop load), using both k-means and $G_i^*$-k means (Fig. 6). As expected, use of the $G_i^*$-k means of the three blocks accentuated groupings but lost individual tree data.

Obviously, the classification based on fruit numbers created clusters of different fruit load, each with lower SD than that of the parent population (Table 3). However, the classification zones based on trunk circumference or canopy size or all three variables were not significantly different in fruit load per tree.

Given a total orchard variability of CV = 78%, the number of samples required for reliable estimation of mean tree fruit load (P = 0.95, PE = 10%) was 235 trees (Eq. 1) or 157 trees when adjusted for the finite sample size (Eq. 2) (Table 1). For this level of reliability, 14, 17 and 111 trees are required (Eq. 1) for sampling of the three classifications based on fruit load, or a total of 142. This represents only a 10% decrease on sample number based on entire orchard variability.
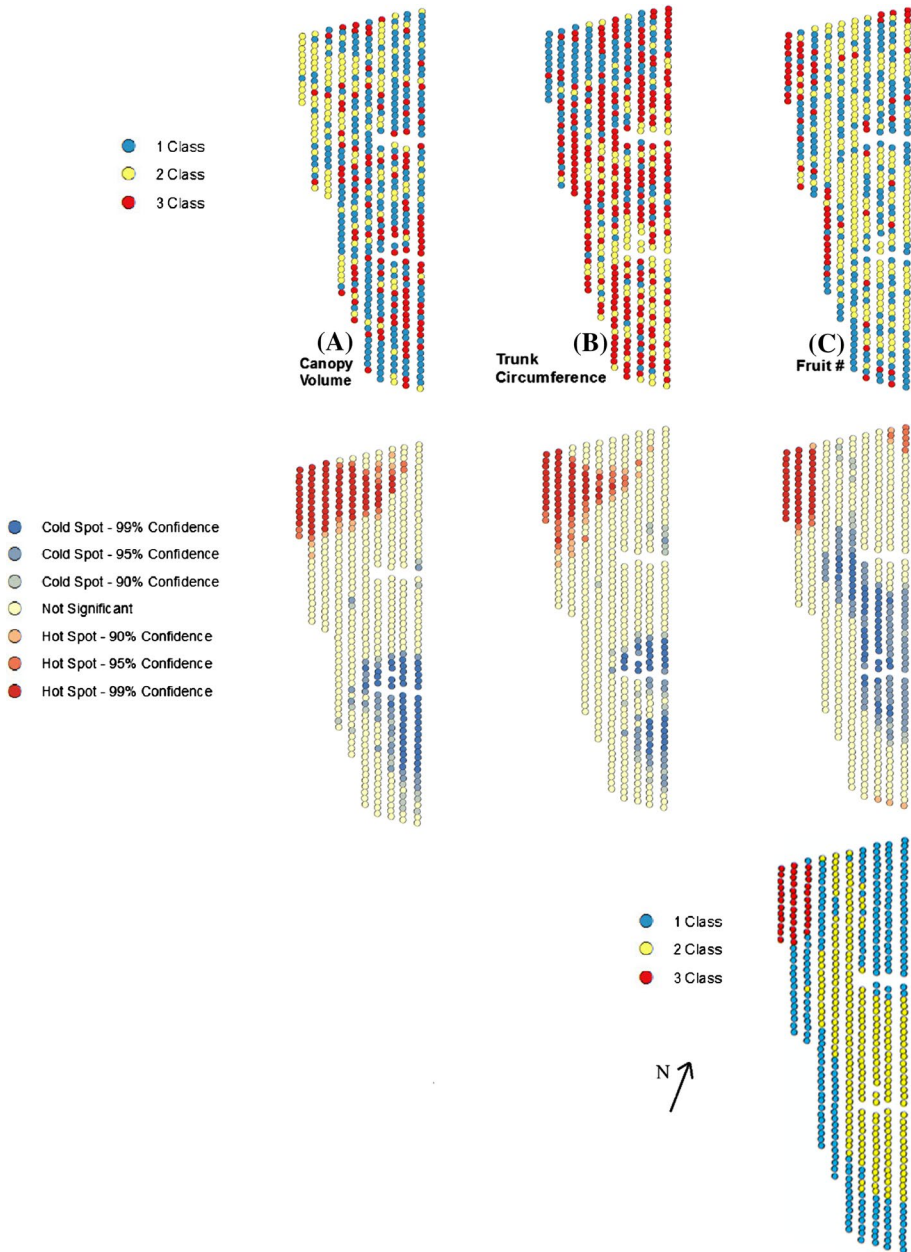
**Fig. 5** Orchard map with trees classified to three clusters by k-means clustering (top row) and $G_i^*$ spatial analysis (middle row) for variables of canopy volume (**a**; cm$^3$), trunk circumference (**b**; cm) and number of fruits per tree (**c**). The Z-score derived from $G_i^*$ analysis on # of fruits per tree was used as the input for k-means clustering (bottom row). In the middle row, colouring of the data points is based on significance level of the cluster classification (Color figure online)
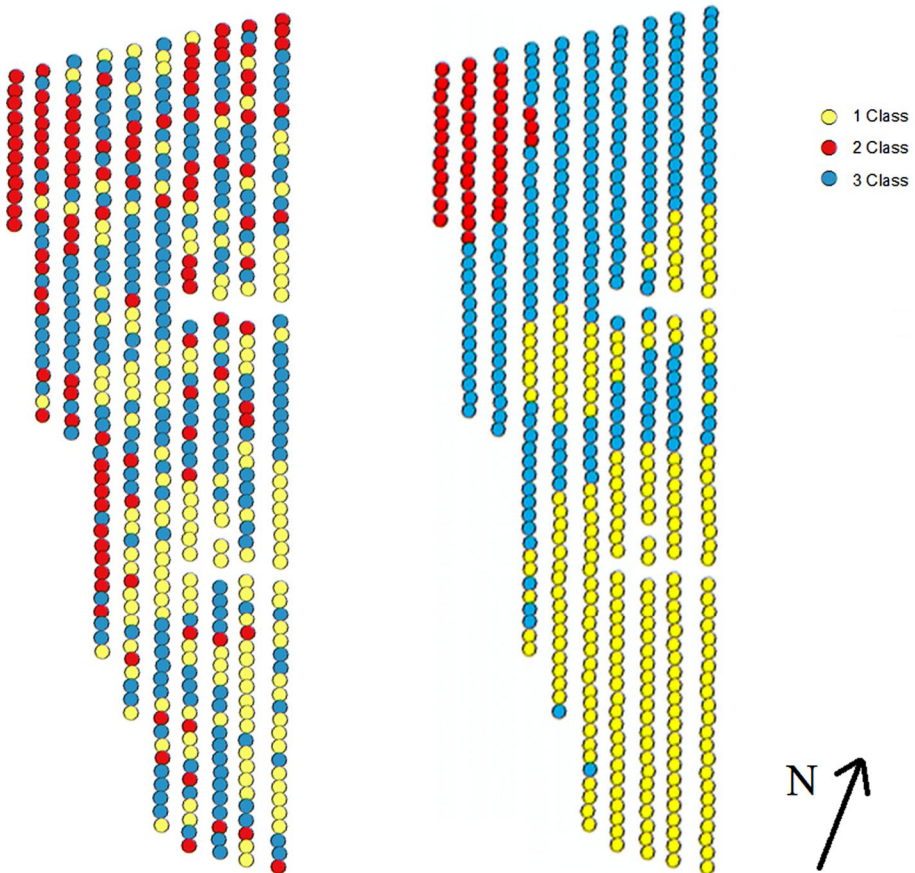
**Fig. 6** Orchard map with trees classified to three clusters by the multi-variates of canopy volume, trunk circumference and fruit number per tree. K-means clustering analysis plots with standardized tree attribute values as input (left) and with input of z-scores from individual $G_i^*$ analysis of attribute levels (right) (Color figure online)

K-means classification on canopy volume could be useful for targeted pruning regimes or for estimation for amounts of plant protection chemical to apply. The use of $G_i^*$ analysis was not useful, masking spatial variation.

## Method comparisons for orchard yield estimates

The high variance observed in mango tree crop yield renders the practice of manual in-field counts impractical in terms of the sample number required, and the commercial best practice count of 5% of trees was demonstrated to deliver an unreliable estimate for all ten orchards considered. Until breeding or agronomic practices deliver decreased tree yield variance, methods are required that allow estimation of orchard yield based on high numbers of trees.

**Table 3** Blocks classified by k-means clustering for the three individual variables and for the combined set of variables, displaying average, standard deviation, # of trees (n) within each class, sample trees required per classification and the adjusted n of sample trees required

| | Average (fruit/tree) | SD (fruit/tree) | n (trees/class) | Sample required (trees) | Adjusted n (trees) |
|---|---|---|---|---|---|
| Fruit # | | | | | |
| Class 3 | 264 | 56 | 65 | 18 | 14 |
| Class 1 | 139 | 31 | 185 | 19 | 17 |
| Class 2 | 37 | 28 | 218 | 225 | 111 |
| Total n | | | | 262 | 142 |
| Trunk circumference | | | | | |
| Class 3 | 69 | 56 | 88 | 257 | 66 |
| Class 1 | 159 | 93 | 138 | 131 | 67 |
| Class 2 | 94 | 77 | 242 | 253 | 124 |
| Total n | | | | 641 | 257 |
| Canopy volume | | | | | |
| Class 3 | 166 | 102 | 98 | 143 | 58 |
| Class 1 | 106 | 78 | 239 | 209 | 111 |
| Class 2 | 71 | 58 | 131 | 261 | 87 |
| Total n | | | | 613 | 256 |
| Combined k-means | | | | | |
| Class 3 | 215 | 73 | 115 | 44 | 32 |
| Class 1 | 81 | 55 | 202 | 177 | 94 |
| Class 2 | 64 | 53 | 152 | 263 | 96 |
| Total n | | | | 484 | 222 |

The total orchard had n=469 trees, with average=110 and SD=89 fruit/tree (multi-view estimated)

**Table 4** Harvest count of fruit and estimation from several methods involving count of a sample of trees or satellite imagery VI and multi-view/dual-view machine vision of the entire orchard (469 trees)

| | Fruit count | % of actual harvest | CV (%) |
|---|---|---|---|
| Actual harvest | 56,720 | 100 | |
| Estimate based on 5% stratified sample (manual count of 23 trees) | 47,955 | 85 | |
| Estimate based on 5% stratified sample (multi-view data) (n=20, 23 trees)) | 53,158 ± 9597 | 94 | 18 |
| Estimate based on random selection of 157 trees (multi-view data) (n=20) | 53,186 ± 3005 | 94 | 6 |
| WV3 VI | 51,944 ± 26,300 | 92 | 45 |
| Dual-view | 51,548 ± 2914 | 91 | 14 |
| Multi-view | 53,520 ± 1960 | 94 | 4 |

Actual harvest is from packhouse record of a strip harvest. Values presented as mean and associated SD when available

The multi-view and dual-view machine vision method and remote sensing method achieved an estimate of orchard yield that was between 91 and 94% of the actual harvest (56,720 fruit), with the multi-view method delivering the most reliable estimate (CV = 4%, Table 4). The high level of uncertainty (CV = 45%) of the satellite VI estimate was due to the high calibration RMSE.

For this orchard, the multi-view method achieved a slope of 1.0043 between machine vision and harvest counts, indicating that effectively all fruit on the canopy were visible from the inter-row. This may not be true for other canopy structures, with fruit hidden from view by the canopy and other fruit. In such cases, accurate crop load estimation would require a 'calibration' of the machine vision count to the actual tree count. Similarly satellite remote sensing estimates are anticipated to require calibration with actual fruit load in each season. However, these remote sensing estimates could be used in classification of trees by fruit number, potentially informing tree selection for such a calibration.

## Conclusion

The variation in tree fruit load between trees in an orchard was high for the ten orchards considered (SD from 34 to 160; CV from 27 to 93%), such that the number of trees required to be counted for a reliable and useful estimate (95% confidence, to 10% error) was prohibitive for manual counts. Attempts to reduce sampling effort by classification of an orchard on the basis of potential fruit load (using surrogates for number of terminals such as trunk circumference or canopy volume) using either k-means or $G_i^*$-k means clustering were unsuccessful, indicative of differences in percentage of terminals that flowered or in fruit retention across the orchard.

For orchards in which canopy architecture allows all fruit to be seen from the inter-row, machine vision technology can allow for estimation of orchard fruit yield without recourse to manual counts. Note that the whole orchard need not be assessed, but rather a sample size consistent with the required precision. For orchards in which fruit are hidden from view, a calibration against manual count would be required. The satellite imagery VI based correlations on fruit yield allow for estimations across large areas, however measurement uncertainty was large, and calibration of the VI—fruit count relationship may be required per orchard or season.

Individual tree yield data should prove useful in precision management programs, and in selection of trees with high harvest indices and low bienniality in bearing.

# References

Esri. (2018). *An overview of the spatial statistics toolbox*. Retrieved September 9, 2018, from http://pro. arcgis.com/en/pro-app/tool-reference/spatial-statistics/an-overview-of-the-spatial-statistics-toolb ox.htm. Redlands, CA, USA.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis, 27*(4), 286–306.

Payne, A. B., Walsh, K. B., Subedi, P. P., & Jarvis, D. (2013). Estimation of mango crop yield using image analysis—Segmentation method. *Computers and Electronics in Agriculture, 91,* 57–64.

Peeters, A., Zude, M., Käthner, J., Ünlü, M., Kanber, R., Hetzroni, A., et al. (2015). Getis–Ord's hot-and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. *Computers and Electronics in Agriculture, 111,* 140–150.

Rasmussen, C., & Williams, C. (2006). *Gaussian processes for machine learning* (pp. 7–13). Massachusetts, USA: The MIT Press.

Robson, A., Rahman, M. M., & Muir, J. (2017). Using worldview satellite imagery to map yield in avocado (*Persea americana*): A case study in Bundaberg, Australia. *Remote Sensing, 9*(12), 1223.

Robson, A. J., Petty, J., Joyce, D. C., Marques, J. R., & Hofman, P. J. (2016). High resolution remote sensing, GIS and Google Earth for avocado fruit quality mapping and tree number auditing. In *Proceedings of the 29th international horticultural congress (IHC2014)* (pp. 589–595). Leuven, Belgium: ISHS.

Stein, M., Bargoti, S., & Underwood, J. (2016). Image based mango fruit detection, localisation and yield estimation using multiple view geometry. *Sensors, 16*(11), 1915.

Taylor, J. A., Praat, J. P., & Bollen, A. F. (2007). Spatial variability of kiwifruit quality in orchards and its implications for sampling and mapping. *HortScience, 42*(2), 246–250.

Thrusfield, M. (1995). *Veterinary epidemiology* (2nd ed.). Oxford, UK: Blackwell Science Ltd.

Yadav, I., Rao, N. S., Reddy, B., Rawal, R., Srinivasan, V., Sujatha, N., et al. (2002). Acreage and production estimation of mango orchards using indian remote sensing (IRS) satellite data. *Scientia Horticulturae, 93*(2), 105–123.

## Affiliations

**N. T. Anderson[1] · J. P. Underwood[2] · M. M. Rahman[3] · A. Robson[3] · K. B. Walsh[1]**

[1]   Institute for Future Farming Systems, Central Queensland University, Building 361 CQU, Bruce Highway, Rockhampton, QLD 4701, Australia

[2]   Australian Centre for Field Robotics, University of Sydney, Sydney, Australia

[3]   Precision Agriculture Group, University of New England, Armidale, Australia