

Estimation of papaya volume and surface area using a dual ellipsoid representation and image analysis

Gabriel Scarmana, Australia

ABSTRACT

Knowledge of volume (V) and surface area (A) of agricultural products can be used in the design of machinery, in predicting amounts of pesticides, and in quantification of bruise, abrasion, and insect damage. A and V are also valuable for calculating rates of post-harvest loss in horticultural produce and monitoring fruit growth

In this study, laser scanning methods and standard hydro-static techniques are in first place used to create accurate and scaled 3D models of a fruit cultivar (i.e. red papaya, *Carica Papaya*) of different dimensions. Based on these models a geometric representation for a rapid estimation of V and A of a given papaya is presented. In essence, this representation relates to a mathematical function based on a 3D shape formed by two halves of ellipsoids that have separate radius dimensions. This representation will be referred to as dual ellipsoid for the remaining of this manuscript.

Statistical data regarding V and A as obtained via the proposed dual ellipsoid is determined by comparing said data with that obtained from the above mentioned hydrostatic and laser scanning techniques. Results show that for the relatively large papaya samples considered, the predicted (via the proposed representation) and the “true” figures for V and A agree within 1.8% and 2.0 % respectively.

The model representation is defined by fruit dimensions which comply with Australian standards for size and volume characteristics of papaya cultivars in general, that is, papaya axial dimensions such as width and length. In this work these axial dimensions are derived from measurements taken on scaled digital images of the fruit using an open source digital image analysis software. Hence, the proposed model may constitute a practical, effective and economic tool for many applications related to field measurement, mapping, growth, monitoring, harvesting, processing, storing and handling of this highly perishable fruit.

In particular it may lead to applications in farm management, production and forecasting. For example, many present practices related to monitoring of fruit growth in the field are still relying on callipers and/or measuring gauges. In this context, a GIS (Geographical Information Systems) field measuring scheme is suggested, capable of storing pictorial information of papayas, including attributes like location (i.e. latitude and longitude) of measurements, weather and time of measurements.

1. INTRODUCTION

In post-harvest processes the physical properties of agricultural products, such as surface area (A), density (D) and mass (M) have wide applications. For instance, one of the critical factors in determining respiratory rate, evaluating colour in separation, heating in heat transfer rate, cooling, and freezing processes is the A . The other crucial component that plays a role in packaging design and storing agricultural products is V (Arshad, 2014).

In addition, the requirement of measuring and/or predicting A and V of fruit items directly or indirectly relate to research studies in all fields of entomology, pathology, physiology, and chemistry. Estimations of A and V also enter schemes of areas of fruit damage, use of pesticides, toxic gases and dusts which are applied in the control of pests (Petros et al., 2014).

V of fruit items can be determined via hydrostatic (i.e., water displacement and/or suspension) techniques. Although accurate and reliable, standard hydrostatic techniques are time consuming and impractical under field conditions. Likewise, laser scanning methods for defining V and A entail the processing of thousands of sample 3D points (many of them redundant) over the fruit surface, requiring precise surface thinning and matching techniques (Dyn et al., 2008), not to mention the economics behind the implementation of adequate laser scanning instrumentation (Ebrahim, 2014).

By way of example, the reader is referred to the work of Villardon et al. (2020). In this article, laser scanning based techniques, including their accuracy and consistency, were applied for determining the V and A of sweet potatoes. Measurements of sweet potato shape features were comparable with currently accepted manual measuring methods, with the added benefits of increased throughput and a non-destructive sampling procedure. In summary, scanner-based A and V computations overcome perceived limitations of simply measuring axial dimensions of fruit items and other similar agricultural products. However and as previously pointed out, they do not constitute a practical measuring method for field applications.

Image processing has also been utilised for determining V and A of fruit items with various degrees of accuracy results (Blasco et al., 2012, Mieszkalski et al. 2017, and Sabliov et al. 2002). In a typical image processing method a whole fruit item is placed on a rotating disk or in a static position and 2D images, rectified for distortions and differing by pre-established or random small angles, are taken to the fruit item (Kim et al. 2021). The contours or profiles of the fruit are extracted from the images and used to reconstruct the 3D shape of the fruit using standard mathematical models.

The V and A calculated by way of these methods require (not often mentioned) the rectification and proper scaling of the object within the image, especially if there are large variations in the size of the object photographed (i.e., small or large papayas). In other words, while the location of the camera and the object remains the same, the distance between the camera and the object is reduced in line with the increasing size of the object of interest (Elder, 2019). Hence, by simply placing a visible scale bar on the background next to the

object and using it for measuring and scaling purposes will inevitably lead to incorrect and/or distorted measurements.

Also, images taken with cameras suffer from lens distortions as it is the case of barrel distortion, and pincushion distortion. As already inferred, these distortions have the effect of misrepresenting the dimensions derived from images unless cameras are calibrated and adequate rectifications are taken into consideration and applied accordingly (Kim et al., 2021).

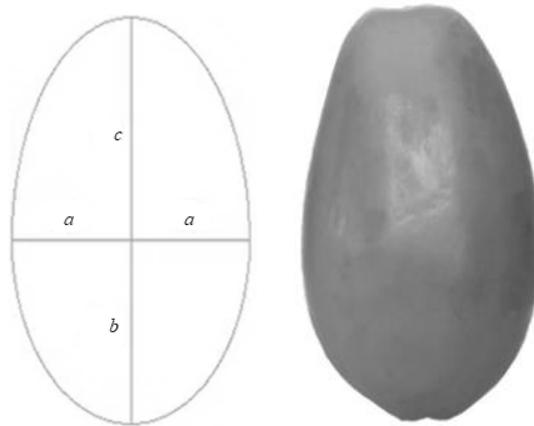
In light of these considerations, the proposed method of determining the A and V of a papaya relates to mathematical equations that are simple to program and sufficiently accurate, and that can predict the A and V based on fewer measurements as possible, that is, the papaya axial dimensions.

Due to the reliability of a hydrostatic technique referred to as the suspension method and laser scanning techniques, they were considered here as a way of establishing a point of truth for accuracy assessment purposes, and in this manner prove statistically the validity and/or the accuracy attainable with the proposed dual ellipsoid model.

2. THE DUAL ELLIPSOID MODEL

From a deterministic viewpoint A and V of papayas can be estimated upon the assumption that they tolerate a similarity to conventional axis-symmetric geometric shapes (i.e., an ellipsoid) from which A and V can be mathematically approximated using measurements taken on the fruit as shown in Figure 1. The computation of A and V based on an ellipsoid is a valid postulation in the case of products like eggs, lemons, limes, oranges, tomatoes (Li et al., 2011) which to a certain extent have consistent axis-symmetric characteristics and favourable ellipse-like properties (Sitkei, 1986).

While papaya may be also considered as having an axis-symmetric geometry, their shape in general is far from directly adhering to a standard ellipsoid per se. However, the shape of a papaya could be associated to a dual ellipsoid formed by two halves of ellipsoids that have separate radius dimensions. One half has the radii of a , a , and b , while the second has the radii of a , a , and c . Hence, the V and A of a papaya can be calculated via the two equation in Figure 1 (Ayres and Mendelson, 2008).



$$V = \frac{2\pi}{3} a^2 (b + c) \quad A = 2\pi a^2 + \pi a \left(\frac{b^2}{\sqrt{b^2 - a^2}} \cos^{-1} \left(\frac{a}{b} \right) + \frac{c^2}{\sqrt{c^2 - a^2}} \cos^{-1} \left(\frac{a}{c} \right) \right)$$

Figure 1 - Dual ellipsoid representation of a section of a papaya and its dimensions in terms of the parameters a , b and c . V is the volume and A is the surface area.

Clearly, these axis-symmetric models do not take into consideration geometric irregularities over the fruit surface. For instance, measurements of width (i.e., twice the parameter a) taken at different locations may differ significantly, thus resulting A and V discrepancies when using the equations in Figure 1. One way of attenuating the impact of these possible differing width measurements can be achieved by way of averaging a number of them using a calliper. Alternatively, the measurement of the equatorial circumference may be adopted for calculating a . In this case a will be the result of averaging all possible a values as derived from the circumference itself.

Diameter tapes have been considered with more consistency than callipers when measuring fruit items considered axis-symmetric because they measure an average of all diameters in all directions (Burkhart and Harold, 1993). Calliper arms only measure one diameter at the time, but since the cross sections of papayas are not completely circular in essence, different diameter measurements are possible. At any rate, practical recommendations for diameter measurements are: (i) the largest and smallest diameter of the section for visible elliptical sections; and for close-to-circular sections (ii) the largest diameter and another perpendicular to the former; or (iii) the diameter of two perpendicular axes taken at random.

In these three cases, the two diameter measurements can be either averaged using the arithmetic mean, or averaged by the geometric mean e.g. for highly elliptical papaya cross sections. Following Cauchy's theorem (1841) it can be argued that the average from a number of random diameter measures using a calliper is comparable to the diameter value obtained from a perimeter measurement with a tape (Rodriguez et al., 2015). In other words, both tools provide analogous results (Biging & Wensel, 1988).

The decision to measure diameter with callipers or circumference by tape often depends on the available tools and resources, tradition and the level of acceptable error (Barack, 2001). An interesting manner of using these measurements and optimise results may be found in the work of Ziaratban et al. (2017) which describes the modelling of V and A area of apples from their geometric characteristics (that is, measurements taken on the fruit) and artificial neural network.

3. POINTS OF TRUTH

The true values for papaya V were computed by measuring the V of small objects based on the Archimedes principle referred to as the suspension method. It involves suspending an object (i.e. a papaya) in a water-filled container placed on an electronic scale. The principle is based on the fact that said object can simply be suspended in water and the change in weight is translated directly into its volume. The reader is referred to Hughes (2005) for a thorough explanation of this method of volume determination, including its advantages when compared to the standard water displacement method.

On the other hand, the true value of A was based on a laser scanning process. The scanner available to the author was the Matter and Form 3D Scanner (<https://matterandform.net/scanner>). The scanning and computation of A for each papaya was time consuming and it was accomplished with a density of approximately 8-10 XYZ points per cm^2 . These points were used to generate triangulated irregular networks or TIN models.

The same points were also used to determine V values for each papaya. Upon comparing the results of V obtained by the above-mentioned suspension method and the V obtained by the scanning process it was found that overall they only differ by 0.46%. This gave the author assurance regarding the dependability of the adopted point of truth values for V and to a confident extent the values of A .

The scanning process was accomplished in approximately 7-8 minutes per individual papaya with the scanner set on the QuickScan mode. However, the time needed for the actual computation of A was in the vicinity of 5 minutes per individual papaya. This involved the so-called thinning or elimination of duplicate and/or redundant overlapping points of the dense cloud of points produced by the scanner. Accuracy of the scanner as claimed by the manufacturer is within $\pm 0.1\text{mm}$ and the maximum object size should be less than 25 cm. (height) and 18 cm (diameter). Maximum weight of the object should not exceed 3 Kg. The scanner operates between temperatures of 15°C and 32°C . The few papayas that exceeded the above dimensions had to be scanned in two parts by sectioning transversally the fruit into two pieces.

4. MATERIALS AND METHODS

Thirty-five (35) consumer grade fruit items from the red papaya cultivar were selected for the tests. All papayas originated from farms located in North Queensland, Australia. Selection was also based on papayas showing relative irregular shapes and size, except for very

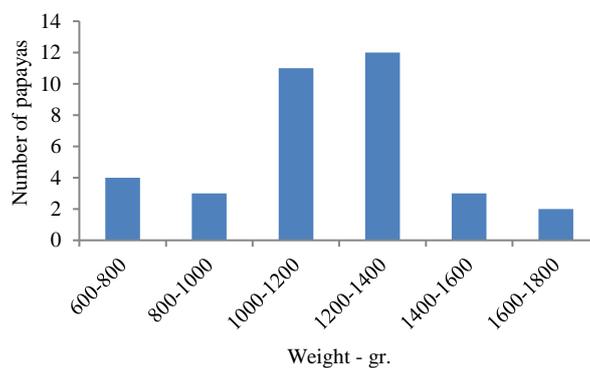
unusually or markedly asymmetrical fruit, but otherwise random samples were used. Figure 2 gives examples of the selected papayas.



Figure 2 - Typical examples of selected papayas

Fruit weight (W) of the samples ranged between 650 gr. to 1800 gr. They were weighed to the nearest 0.5 gr. Fruit V was determined by the water suspension method as described in section 3. Figure 3 shows the frequency distribution of W and V of the fruits. The strong correlation existing between V and W is graphically and analytically shown in Figure 4.

For the sample papayas considered here this assumption was verified via regression analysis using the Trend tool of Microsoft Excel (c). Figure 4 shows the plot of papaya W versus their corresponding V , including the linear function that estimates V from W . A perfect coefficient of determination $R^2 = 1$ was not expected as two papayas can have the same V but different W which can be attributed to the shape of the papaya and/or the V of the internal seeds' cavity.



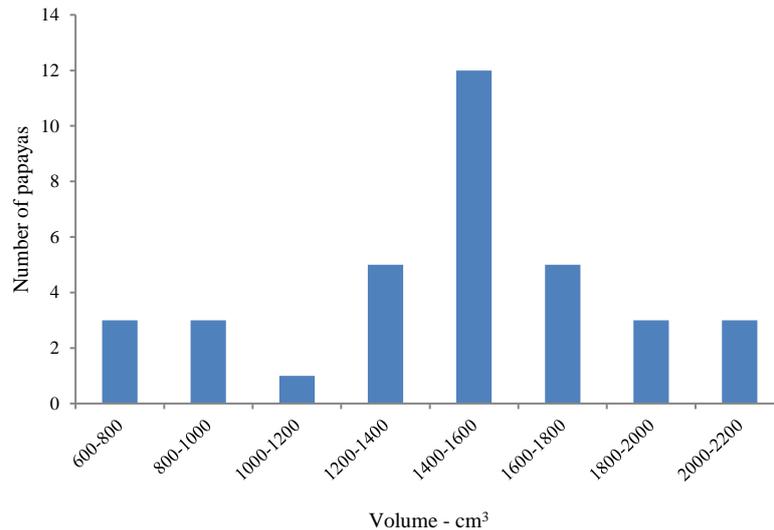


Figure 3 - Frequency distributions of the 35 papayas V and W processed in this work.

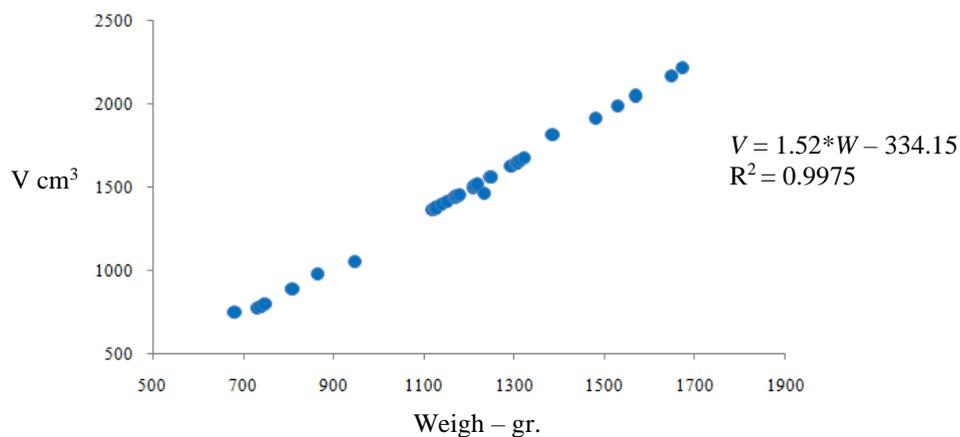


Figure 4 - Linear model for V (cm^3) for the total observations based on papaya W (gr.)

Actual papaya A was estimated by scanning each papaya using the technique mentioned in section 3. The accuracy of the technique was further verified by scanning a near perfect sphere of known diameter (a solid fibreglass ball, 110 mm. diameter) in the same way. Estimation of A for the sphere by the scanning method closely approximated (0.35% difference) the value obtained by mathematical calculation based on the sphere diameter.

Predicted values of V and A using the proposed dual ellipsoid model were then obtained for each papaya using the average of the major and minor diameters of the fruit item, and its length L . These dimensions were measured with an off-the shelf large vernier calliper (300mm, +/- 0.05 mm. accuracy as claimed by manufacturer). Statistical figures related to the

accuracy of the equations in Figure 1 were established upon comparing the results of the point of truth of V and A and the predicted values obtained from the equations of the dual ellipsoid representation. This is further illustrated in the ensuing section.

To determine the variables b and c in Figure 1 and thus estimate the values of A and V it was thought of defining a standard ratio between the length L of the papaya and the point of maximum width. This ratio was used to simplify the calculations of the long and short polar radii b and c .

The typical ratio for the thirty-five papayas was determined by averaging all ratios and considering the principles given in the paper of White and Alspach (1996). In this paper, digitised video images of fruits (i.e., pears) were taken and three different height and width ratios describing fruit shape were estimated from measurements on a large population 17 pear cultivars.

Based on this initiative, Figure 5 illustrates the definition of the standard ratio r employed in this work and applied for all 35 papayas. The graph in this same figure illustrates the fluctuation of the 35 values of the ratios and the resulting average value. This ratio r was determined to be equal to 0.395 with a standard deviation of 0.052 with maximum and minimum values ranging between 0.425 and 0.384 respectively.

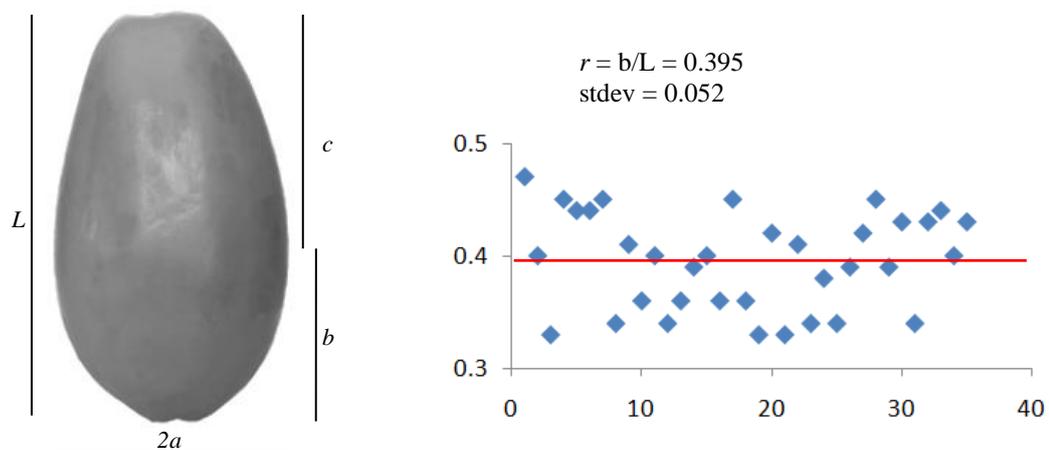


Figure 5 - Ratio relating the length L of a papaya to the parameter b and c used in estimating V and A via the dual ellipsoid model. Right, is the graph showing all ratios r for the 35 papayas. The red line represents the average of all 35 ratios. In this instance $r = 0.395$.

5. RESULTS

Table 1 illustrates the RMSE (Root Mean Squared Error) of the differences between the point of truth values of V and A and the predicted values using the proposed equations. The differences in terms of the overall percentage are also included. The table also indicates the maximum and minimum differences from the “point of truth” values.

	RMSE	Overall Diff. %	Max. Diff.	Min. Diff.
A	+/- 2.03 cm ²	2.00	3.75 cm ²	-4.33 cm ²
V	+/- 2.05 cm ³	1.82	4.00 cm ³	-5.03 cm ³

Table 1 - RMSE of the differences between predicted values of A and V using the dual ellipsoid model and those determined by the point of truth methods. The overall difference in percentage and maximum and minimum deviation from the point of truth values of A and V are also shown.

6. A FIELD MEASURING SYSTEM

The decisions on plantation management are in most cases based on visual and subjective observations, in other words a result of experience rather than based on objective information (Manfrini et al. 2012). Standard monitoring strategies are mostly used to release information on environmental variation, which may not directly/easily relate to crop status, thus making at times the interpretation of the seasonal production scheme not easy.

In papaya production schemes the monitoring of fruit size can relate to papaya quality (Nantawan et al., 2017). However, fruit growers need monitoring schemes that are both rapid and sufficiently truthful (Nissen, 2018). Hence, the dimensional measuring system presented in this section may help fill this knowledge gap. The intention is to provide a support system at critical times of the crop season, and to provide fruit dimensional information during the fruit growing stages to help better plan harvest and associated post-harvest logistics.

In this context, measurements on a papaya growing stages on trees could be taken at intervals during the growing season. These measurements could be recorded using a mobile phone and a square ruler as illustrated in Figure 5. This simplistic method allows a single person to record a scaled image of a papaya in a very short period of time, including GNSS (Global Navigation Satellite System) coordinates for location and mapping purposes. Any observation and or attribute of interest can also be recorded if standard and/or mobile field data collection systems like ESRI Collector (<https://esriaustralia.com.au/arcgis-collector>) or Avenza Map (<https://www.avenzamaps.com/>) are considered.

For the purpose of obtaining accurate dimensional data from a single photo, an open-source image analysis software referred to as ADI (Analysing Digital Images) was considered. ADI can be used for analysing each scaled image captured in the field and thus derive the parameters L , a and c of the fruit. ADI is educational image analysis software available on either PCs or Macs. ADI is a software adapted from Real Basic to Java with additional work

on the Java version done at the University of Massachusetts between 2013 and 2015.

In a papaya plantation environment scaled images of the fruit can be based on a tree selection strategy (Manfrini et al., 2015) for the purpose of sampling and data collection. The trees and rows selection can be done randomly, and on each tree an established number of images of different fruit samples are taken as shown in Figure 6.

The image in this figure was taken using a current android mobile phone. The papaya was imaged on its resting position on the tree. A square ruler is also visible. Images with an overall spatial resolution of 2560x1920 pixels were stored as PNG (Portable Network Graphics) files.

Pattern analysis of imaged objects requires all images be scaled to a uniform number of pixels per unit length. Photographs were taken vertical to the fruit (straight on) thus minimizing perspective distortion. The location of the square ruler was crucial for obtaining consistent measurements in terms of accuracy. The square ruler was placed by hand as much as possible levelled with the papaya middle section so to be aligned with the plane or maximum diameter. In other words, the square ruler was set to correspond with the same plane as the papaya longitudinal section (so that the scale and the profile of the papaya could be viewed similarly by the camera).

The phone recorded the location where the photograph was taken, the papaya ID, whereas the fruit dimensional parameters were determined in the office using the ADI software. All measurements taken on the image via the ADI software can be stored and exported to program for data analysis and documentation (i.e., Microsoft Excel). In this instance the L of the papaya was measured to be 23.3 cm, with c and b equal to 16.55 cm. and 6.73 cm. respectively. Hence, the value of a was equal to 3.25 cm.

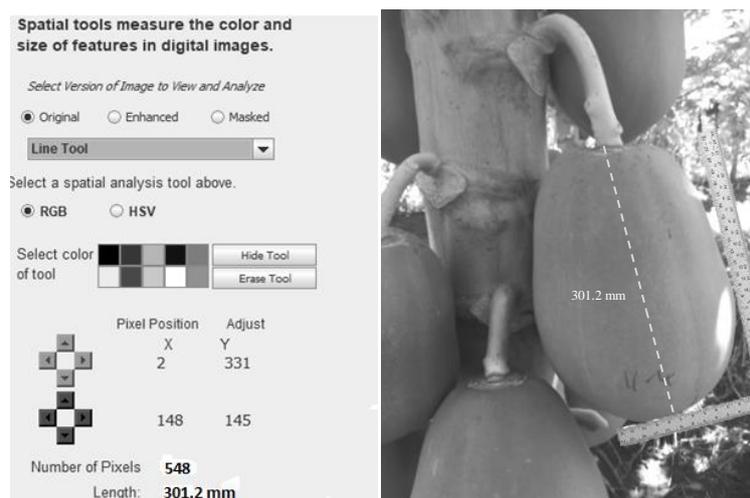


Figure 6 - The field measuring system captures the papaya dimensional data on a tree. A square ruler is visible in the image for scaling/dimensioning purposes. The ADI software is used to obtain accurate measurements of the length (301.2 mm) and width (not shown) of the papaya still attached to the tree.

7. CONCLUSIONS AND DISCUSSION

This study relates to a model that allows a mathematical description of the geometry of papayas with only a few parameters. The dual ellipsoid model approach described here provides a simple and practical tool for characterizing important papaya physical dimensions such as V and A .

The study findings confirmed the capability and accuracy of the model in predicting V and A of papaya fruit. The proposed equations are easy to program using conventional software such as Microsoft Excel (c).

Clearly, methods of measuring V and A by water suspension and laser scanning techniques would yield more accurate results but these methods are time consuming, they require measuring equipment and associated software and are impractical for field measurements (i.e., assessing papayas still attached to trees).

The following conclusions can be drawn from this study:

- The adopted dual ellipsoid model can define the geometry of papayas in terms of minimal dimensional parameters and shape coefficients (i.e., axial dimensions).
- The proposed papaya dual ellipsoid model was validated satisfactorily in terms of precision and accuracy when compared to actual papaya geometry.
- The model has a potential for application in fruit-related computer aided design and simulations and may suggest ideas for further geometric modelling of agricultural products.
- Implementation of the proposed model is scalable can be adjusted or modified so to be used in fruit growth monitoring in the field as needed.

Further studies are required to determine whether the potential transformations of the dual ellipsoid model may be suitable for other cultivars different than the one examined here but having similar shapes (i.e., avocados or pears). Hence, departing from this model, it may be possible to design new geometries applicable to other papaya cultivars (i.e., yellow papayas). Finally, all papayas included in the above tests were consumed and none of them were disposed or destroyed.

REFERENCES

Arshad, M. (2014) "Significance of physical properties of apple fruit influenced by pre-harvest orchard management factors". *European Journal of Agriculture and Biology*. 4. 82-89.

Ayres F. and Mendelson E. (2008) *Calculus. Schaum's Outlines*. McGraw-Hill Professional. pp. 244–248. ISBN 978-0-07-150861-2, 244 pages.

Barack C. (2001) "Forest Measurement and Modeling - Measuring trees, stands and forests for effective forest management". Computer-based course resources for Forest Measurement and Modeling (FSTY2009) at the Australian National University.

Biging G., Wensel L. (1988) "The Effect of Eccentricity on the Estimation of Basal Area and Basal Area Increment of Coniferous Trees". DOI:10.1093/FORRESTSCIENCE/34.3.621
Corpus ID: 89112416. Environmental Science, Mathematics. Forest Science.

Blasco J., Aleixos N., Gutierrez A. and Moltó E. (2012) "Measurement of Physical Properties of Fruits Using Image Analysis". In Physical Properties of Foods - Novel Measurement Techniques and Applications (Chapter 11). Editor Ignacio Arana. CRC Press. 422 pages.

Burkhart A. and Harold E. (1993) Forest Measurements. Published by McGraw-Hill College, 1993 ISBN 10: 0070025568 ISBN 13: 978007002556

Dyn, Nira & Iske, Armin & Wendland, Holger. (2008). "Meshfree Thinning of 3D Point Clouds. Foundations of Computational Mathematics. 8.409-425.10.1007/s10208-007-9008-7.
Ebrahim M. (2014) "3D Laser Scanners: History, Applications, and Future". 10.13140/2.1.3331.3284.

Elder J. (2009) "Photographic reference scale". smallpond.ca/jim/scale. Ottawa, Canada.

Hughes S. W. (2005) "Archimedes revisited: A faster, better, cheaper method of accurately measuring the volume of small objects". In Physics Education, Vol. 40, Number 5 · DOI: 10.1088/0031-9120/40/5/008.

Kim, J., Bae, H. and Lee S. (2021) "Image Distortion and Rectification Calibration Algorithms and Validation Technique for a Stereo Camera". *Electronics* 2021, 10, 339. doi.org/10.3390/electronics10030339.

Li X., Pan Z., Upadhyaya S. K., Atungulu G. and Delwiche M. (2011) "Three-Dimensional Geometric Modelling of Processing Tomatoes" Transaction of the American Society of Agricultural and Biological Engineers, ASABE. 54(6): 2287-2296.

Manfrini L., Pierpaoli E., Zibordi M., Morandi B., Muzzi E., Losciale P., Corelli Grappadelli L. (2015) "Monitoring Strategies for Precise Production of High Quality Fruit and Yield in Apple in Emilia-Romagna" Chemical Engineering Transactions, 44, 301-306.

Mieszkalski L. and Wojdalski J. (2017). "A mathematical method for modelling the shape of apples, Part 1. Description of the method" International Journal on Economics of Technology and Modelling Processes ECONTECHMOD, Vol. 06, No. 3, 97-104.

Nantawan M., Kanchana C. and Ford R. (2017) "Papaya evaluation handbook". Project Number: PP15000. Horticulture Innovation Australia Limited.

Estimation of Papaya Fruit Volume and Surface Area Using a Dual Ellipsoid Representation and Image Analysis (11346)
Gabriel Scarmana (Australia)

FIG Congress 2022
Volunteering for the future - Geospatial excellence for a better living
Warsaw, Poland, 11-15 September 2022

Nissen R, Bound S., Adhikari R. and Cover I. (2018) “Factors affecting post-harvest management of apples: a guide to optimising quality” ISBN: 978-1925646-16-0. Tasmanian Institute of Agriculture, Fruit Growers Tasmania, and Department of Agriculture and Water Resources.

Petros D., Escudero L. and Ioriatti C. (2015) “Integrated Fruit Production and Pest Management in Europe: The Apple Case Study and How Far We Are From the Original Concept?” In *Insects*; 6(3): 626-657.

Rodriguez-Puerta F. and Blanco R. And Aunós A. (2015) ”Diameter versus girth: Which variable provides the best estimate of the cross-sectional area?. *Forest Systems*. 24. e033. 10.5424/fs/2015243-05945.

Sabliov C. M., Boldor D., Keener K.M. and Farkas B. E. (2002). “Processing Method to Determine Surface Area and Volume of Axi-Symmetric Agricultural Products”, *International Journal of Food Properties*, 5:3, 641-653.

Sitkei G. (1986) ”Physical Properties of Agricultural Materials”. *Data Developments in Agricultural Engineering*, ISSN: 0167-4137, Vol. 8, Issue: C, Page: 13-31.

Villordon A., Gregorie C. J. and LaBonte D. (2020) ”Direct Measurement of Sweetpotato Surface Area and Volume Using a Low-cost 3D Scanner for Identification of Shape Features Related to Processing Product Recovery”. *HortScience*. DOI: 10.21273/HORTSCI14964-20. Volume 55: Issue 5.

White A. and Alspach P. (1996) “Variation in fruit shape in three pear hybrid progenies” *24:4*, 409-413, DOI: 10.1080/01140671.1996.9513978

Ziaratban A., Azadbakht M. and Ghasemnezhad A. (2017) “Modeling of volume and surface area of apple from their geometric characteristics and artificial neural network”. *International Journal of Food Properties*, 20:4, 762-768.

CONTACTS

Dr Gabriel Scarmana
School of Surveying and Built Environment.
University of Southern Queensland, Australia.
Gabriel.Scarmana@usq.edu.au