Advances in Animal Biosciences: Precision Agriculture (ECPA) 2017, (2017), **8:2**, pp 477–480 © The Animal Consortium 2017
doi:10.1017/S2040470017001133

advances in animal biosciences

Orange tree canopy volume estimation by manual and LiDAR-based methods

A. F. Colaço[1†], R. G. Trevisan[1], J. P. Molin[1], J. R. Rosell-Polo[2] and A. Escolà[2]

[1]*University of São Paulo, 'Luiz de Queiroz' College of Agriculture, Biosystems Engineering Department, Piracicaba-SP, Brazil;* [2]*University of Lleida – Agrotecnio Center, School of Agrifood and Forestry Science and Engineering, Department of Agricultural and Forest Engineering, Research Group on AgroICT & Precision Agriculture, Lleida - Catalonia, Spain*

*LiDAR (*Light detection and ranging*) technology is an alternative to current manual methods of canopy geometry estimations in orange trees. The objective of this work was to compare different types of canopy volume estimations of orange trees, some inspired on manual methods and others based on a LiDAR sensor. A point cloud was generated for 25 individual trees using a laser scanning system. The* convex-hull *and the* alpha-shape *surface reconstruction algorithms were tested. LiDAR derived models are able to represent orange trees more accurately than traditional methods. However, results differ significantly from the current manual method. In addition, different 3D modeling algorithms resulted in different canopy volume estimations. Therefore, a new standard method should be developed and established.*

Keywords: tree crops, laser scanner, convex-hull, alpha-shape, citrus

Introduction

Canopy volume is an indicator of growth, health and yield potential in tree crops. It is also used for estimating sparing and fertilizer requirements. The current methods for canopy volume estimation in Brazilian orange groves are quite simplistic. One of the most used considers the tree volume as the volume of a cube enclosing the whole tree. Each side of such a cube is measured manually with a measuring tape. Besides its inaccuracy, the manual method is time consuming and requires a significant number of sampled trees for an adequate representation of the whole grove.

One alternative for the manual method is the ranging sensor (Dworak *et al*., 2011). Such technology permits not only a higher accuracy on the estimation of canopy geometric parameters but also higher amount of collected data. These data can be georeferenced and acquired throughout the entire field in order to characterize spatial variability. Once spatial variability of canopy volume is known, inputs can be applied in a variable-rate approach according to the tree size variation. Such applications can be carried out at the same time as sensor readings are acquired (*i.e* in "real time") (Escolà *et al*., 2013). For these reasons, ranging sensors meet the needs of precision horticulture and represent great potential to enhance management in tree crops.

Two types of ranging sensors are mostly investigated in research, the ultrasonic and the LiDAR sensors. These sensors are usually mounted facing the side of the tree row being able to estimate the distance to the vegetation at different heights along the canopy. LiDAR sensors are considered a better solution than ultrasonic sensors because they are usually more accurate, more rapid, and permit distance measurements in multiple directions from one single sensor.

In the past decade, data acquisition and processing methods were developed based on LiDAR technology aiming at providing 3D models of trees and retrieve canopy volume information (Rosell and Sanz, 2012). Rosell *et al*. (2009) described a mobile terrestrial laser scanner (MTLS) in which raw LiDAR sensor data was transformed into a 3D point cloud representing the laser beam impacts on the canopy. Further works showed how to extract geometric canopy information from 3D point clouds obtained by similar data acquisition systems (Escolà *et al*., 2016). Different methods are available for computing canopy volume from point clouds. Auat Cheein *et al*. (2015) classified such methods into two main approaches, one referred as the occupancy grid, in which small regular sized objects are created inside the point cloud structure; and another based on connecting the outer points of the point cloud in order to reconstruct the shape of the object formed by the point cloud structure. Both approaches were tested by the authors and were considered viable methods for estimating canopy volume.

Orange groves in Brazil cover approximately 500,000 ha of land (CONAB, 2013). Spray applications are crucial operations in the management of the groves and might occur at least once a month during the cropping season. Spraying recommendations

† E-mail: andre.colaco@usp.br

are currently based on manual and poorly accurate measurements of canopy volume. Sensor-based estimation should greatly improve the management of these groves.

The hypothesis of this work is that LiDAR-based methods for canopy geometry estimation should be more accurate than manual methods. However, because different methods and algorithms can be used for processing LiDAR data, different results might be achieved evidencing the need for stablishing new standard methods of canopy volume estimation.

The objective of this work was to compare estimations of orange tree canopy volume by six different methods; two based on traditional manual methods and four based on LiDAR technology using different processing algorithms.

Material and Methods

Data acquisition

A mobile terrestrial laser scanner was developed based on a 2D LMS 200 LiDAR sensor (Sick, Waldkirch, Germany) and an RTK (*Real Time Kinematic*) GR3 GNSS receiver (Topcon, Tokyo, Japan). The system was mounted on an ATV vehicle as shown in Figure 1. The LiDAR sensor collected distance values in 181 directions every 13.3 ms (75 Hz) at one-degree angular steps along a vertical plane of the tree canopy. The vehicle moved along the two sides of the tree row at 3.3 m s^{-1}.

A commercial orange grove located in the state of São Paulo, Brazil, was scanned using the developed system. 25 trees inside the grove were selected for this study. The variety of the trees was "Valencia" grafted to "Swingle" rootstock. Trees were six years old at the time of data acquisition (October of 2015).

Data processing

The first step of data processing consisted in creating a georeferenced 3D point cloud representing the laser beam impacts over the canopy. Del-Moral-Martínez *et al*. (2015) give details about this process. The following step is to exclude points that did not represent the target tree canopy (soil and neighboring plants). This step was carried out by visually identifying such points and using tools for selecting and deleting points available in the CloudCompare 2.6.2 software. With the final point cloud from each of the 25 selected trees, the next step was to compute the canopy volume. Two types of surface reconstruction algorithms were applied, the *convex-hull* and the *alpha-shape* using the R software packages *grDevices* and the *alphashape3d*, respectively. Both algorithms connect the outer points of the cloud in order to produce the surface of the enveloping object. Unlike the *convex-hull*, the *alpha-shape* permits the creation of concave objects. The level of concavity is given by setting the index α (lower index results in greater concavities). Three indexes were tested: 0.25, 0.50 and 0.75.

Two types of canopy volume estimation inspired on traditional manual methods were also tested. The first one, which is the most common among Brazilian orange growers, consisted in measuring the volume of a cube that encloses the entire tree. This method is hereafter referred as the cube-fit. The second, referred as the cylinder-fit, considered the canopy volume as two-thirds the volume of a cylinder that encloses the tree. The dimensions of the cube and cylinder were not measured manually as it would be normal. This task usually produces uncertain measurements due to the difficulty of visually defining the limits of the canopy. Besides, it is highly dependent on the person responsible for the task, meaning that usually a systematic error is produced. In order to minimize such errors, the figure dimensions were collected directly from the point cloud using distance-measuring tools available in the CloudCompare 2.6.2 software. Each dimension was given by the average of three measurements.

Results and discussion

After the LiDAR scanning and the point cloud generation, the average number of points representing each tree was 12,100 distributed into approximately 60 perpendicular transects. The 3D models from the proposed algorithms applied to a single tree are shown in Figure 2. It is noticeable how the *convex-hull* model apparently produced a larger object, which occurs because salient branches enlarges the hull structure.

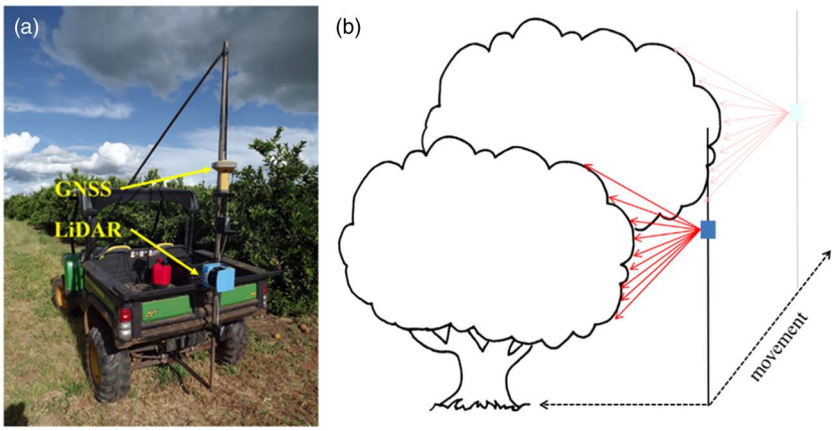

Figure 1 LiDAR sensor and GNSS receiver mounted on an ATV vehicle (a); LiDAR sensor (in blue) facing the side of the tree row (b).

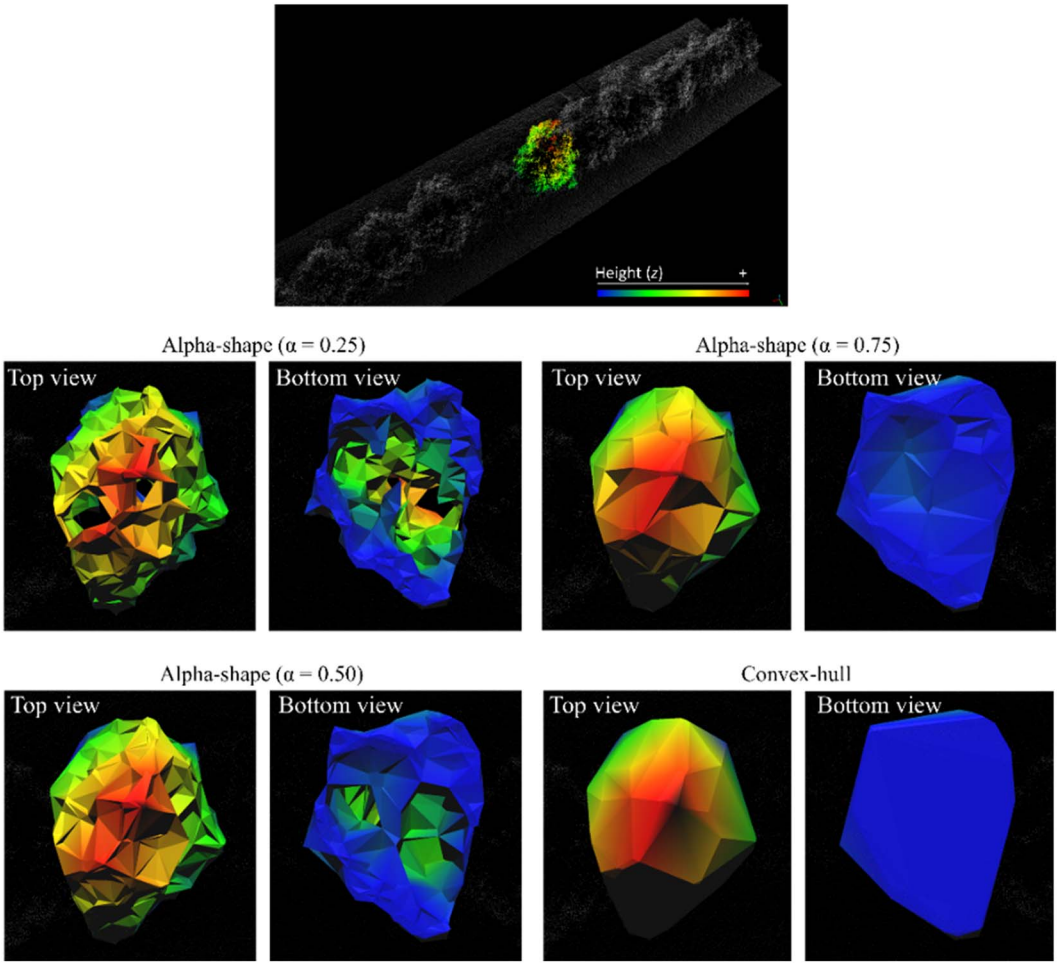

Figure 2 3D canopy structure of a single tree modeled by different algorithms. *Top*: point cloud section and selected single tree. *Middle*: Alpha shape models (left, $\alpha = 0.25$; right, $\alpha = 0.75$). *Bottom left*: Alpha shape model, $\alpha = 0.50$. *Bottom right*: Convex-hull model.

As expected, concavities of the canopy were better represented by the *alpha-shape* algorithm. Because this algorithm permits the representation of concave structures, it is reasonable to consider it a suitable model for representing the tree canopy. However, the index α should be appropriately set (not too high neither too low). A low index might produce disconnected structures forming holes inside the canopy, which is not desirable (see *alpha-shapes* with α set to 0.25 and 0.50 in Figure 2). A guideline for setting α is to choose a value that produces the smallest volume while keeping the canopy as a whole. That result was obtained when α was set to 0.75.

The volumes for each of the 25 individual trees calculated by the proposed algorithms and by the manual methods are showed in Figure 3. The plant IDs were given based on a rank from the smallest to the largest tree according to the *alpha-shape* ($\alpha = 0.75$) and *convex-hull* computation. It is to be noticed that this ranking is roughly followed by all the other algorithms. Generally, the manual methods produced a ranking with some disagreement with the other algorithms, which indicated a certain level of randomness in the manual measurements. The rough simplification of the canopy structure by the manual method based on the cube-fit overestimated the canopy volume in relation to all the other algorithms. The measurements from the cylinder-fit method were mainly close to the ones computed with the alpha-shape with index of 0.50.

Table 1 shows the descriptive statistics of the measurements of the 25 orange trees. A noticeable difference in all statistical parameters between different computation methods is evident. As expected, the *convex-hull* model resulted in the highest mean canopy volume among the proposed algorithms, followed by the *alpha-shape* with higher index α. The volume retrieved from the *alpha-shape* with α set to 0.25 was significantly smaller than that of other algorithms, due to disconnected structures formed by the model and significant voids inside the canopy, as viewed in Figure 2. On the other end, the cube-fit method resulted in a significantly larger volume than all evaluated methods. Regarding the cylinder-fit, because it considered only two thirds of the cylinder, the final canopy volume is not as large as in the cube-fit method. The cylinder-fit method got closer results to the *alpha-shape* with index of 0.5, which, as noticed in Figure 1, produced voids inside the 3D model. The differences pointed out between canopy volume estimations indicate the need to adopt some new standard method.

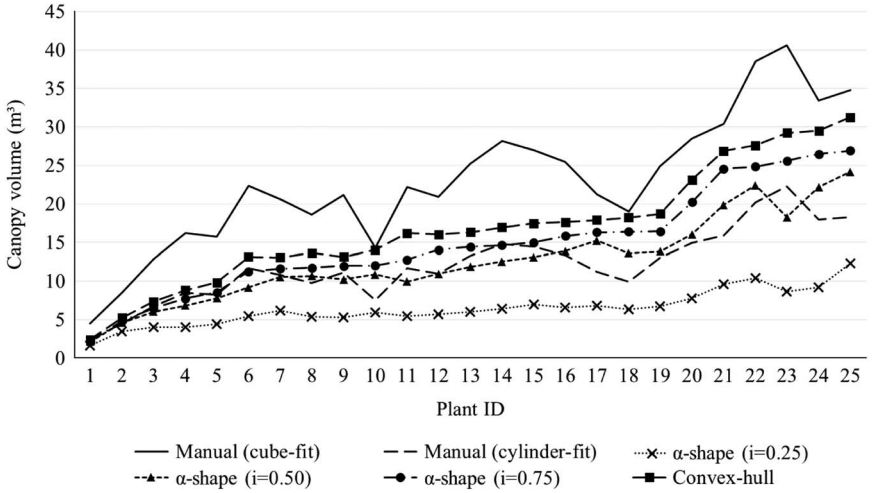

Figure 3 Canopy volume of 25 individual trees computed by different methods. Plant IDs were given based on a rank from the smallest to the largest tree according to the *alpha-shape* ($\alpha = 0.75$) and *convex-hull* computation.

Table 1 *Descriptive statistics of canopy volume of 25 individual trees by different methods*

	Mean	St. Dev.	Min.	Max.	Range	Coef. Var.
			m^3			%
α-shape (i = 0.25)	6.43	2.26	1.65	12.28	10.63	35.11
α-shape (i = 0.50)	12.66	5.44	2.12	24.15	22.03	42.98
α-shape (i = 0.75)	14.92	6.64	2.22	26.89	24.68	44.48
Convex-hull	16.95	7.46	2.42	31.26	28.84	44.04
Manual (cylinder-fit)	12.15	4.59	2.35	22.31	19.96	37.80
Manual (cube-fit)	23.01	8.54	4.48	40.56	36.08	37.14

It is also noticeable from Table 1 that the manual methods presented lower coefficient of variation than the 3D modeling algorithms, with exception to the *alpha-shape* with index 0.25. That might occur because the cube-fit and cylinder-fit represent the tree using less complex figures which are less variable and often do not represent the actual volume of the canopy, *i. e.*, trees with different volume in reality might present similar volumes when these methods are used.

Conclusions

LiDAR-based methods provided more realistic representation of the canopies than current manual methods and should be considered as a new standard for canopy volume estimations.

Important differences among the evaluated methods of canopy volume estimation in orange trees were found. The average canopy volume retrieved by the cube-fit method (the current practice in Brazilian orange groves) was approximately 150% larger than the volume from the alpha-shape with 0.75 index. The alpha-shape modeling with α set to 0.75 was considered an adequate method since it permitted the representation of concave structures of the canopy without creating disconnected objects or voids inside the canopy.

Methods based on cylinder or cube objects were less capable of representing tree size variability than LiDAR-based ones.

Testing more 3D modeling options in a larger amount of sampled trees is still needed. Nevertheless, some guidelines and promising results were presented in this study.

References

Auat Cheein FA, Guivant J, Sanz R, Escolà A, Yandún F, Torres-Torriti M and Rosell-Polo JR 2015. Real-time approaches for characterization of fully and partially scanned canopies in groves. Computers and Electronics in Agriculture 118, 361–371.

CONAB. COMPANHIA NACIONAL DE ABASTESCIMENTO 2013. Acompanhamento de safra brasileira: laranja, terceiro levantamento, dezembro/2013. Brasília, 16p.

Del-Moral-Martínez I, Arnó J, Escolà A, Sanz R, Masip-Vilalta J, Company-Mesa J and Rosell-Polo JR 2015. Georeferenced Scanning System to Estimate the Leaf Wall Area in Tree Crops. Sensors 15 (4), 8382–8405.

Dworak V, Selbeck J and Ehlert D 2011. Ranging sensors for vehicle-based measurement of crop stand and orchard parameters: a review. Transactions Of The Asabe 54 (4), 1497–1510.

Escolà A, Martínez-Casasnovas JA, Rufat J, Arnó J, Arbonés A, Sebé F, Pascual M, Gregorio E and Rosell-Polo JR 2016. Mobile terrestrial laser scanner applications in precision fruticulture/horticulture and tools to extract information from canopy point clouds. Precision Agriculture, in press. DOI: 10.1007/s11119-016-9474-5

Escolà A, Rosell-Polo JR, Planas S, Gil E, Pomar J, Camp F and Solanelles F 2013. Variable rate sprayer. Part 1 – Orchard prototype: Design, implementation and validation. Computers and Electronics in Agriculture 95, 122–135.

Rosell-Polo JR, Llorens J, Sanz R, Arnó J, Ribes-Dasi M, Masip J, Escolà A, Camp F, Solanelles F, Gràcia F, Gil E, Val L, Planas S and Palacín J 2009. Obtaining the three-dimensional structure of tree orchards from remote 2D terrestrial LIDAR scanning. Agricultural and Forest Meteorology 149 (9), 1505–1515.

Rosell-Polo JR and Sanz R 2012. A review of methods and applications of the geometric characterization of tree crops in agricultural activities. Computers and Electronics in Agriculture 81, 124–141.

Advances in Animal Biosciences: Precision Agriculture (ECPA) 2017, (2017), **8:2**, pp 481–486 © The Animal Consortium 2017
doi:10.1017/S2040470017000449

advances in animal biosciences

Mapping properties of an asynchronous crop: the example of time interval between flowering and maturity of banana

J. Lamour[1,2†], O. Naud[1], M. Lechaudel[3] and B. Tisseyre[1]

[1]*UMR ITAP, Irstea - Montpellier SupAgro, France;* [2]*Compagnie Fruitière, Marseille, France;* [3]*CIRAD,UMR Qualisud, F-97130 Capesterre Belle-eau Guadeloupe, France*

Precision agriculture for banana crops has been little investigated so far. The main difficulty to implement precision agriculture methods lies in the asynchronicity of this crop: after a few cycles, each plant has its own development stage in the field. Indeed, maps of agronomical interest are difficult to produce from plant responses without implementing new methods. The present study explores the feasibility to derive a spatially relevant indicator from the date of flowering and the date of maturity (time to harvest). The time between these dates (TFM) may give insight in spatial distribution of vigor. The study was carried out using production data from 2015 acquired in a farm from Cameroon. Data from individual plants that flowered at different weeks were gathered so as to increase the density of TFM sampling. The temporal variability of TFM, which is induced by weather and operational constraints, was compensated by centering TFM data on their medians (TFMc). The mapping of TFMc was obtained using a classical kriging method. Spatial structures highlighted by TFMc either at the farm level or at the plot level, suggest that such maps could be used to support agronomic decisions.

Keywords: banana, mapping, asynchronous, maturity

Introduction

Precision agriculture for banana crops has been little investigated. However, this production is of major economical importance, with more than 100 million tons produced in 2013 (FAO, 2013). Compared to crops such as wheat, corn, canola, etc. whose mapping of production properties is common in precision agriculture, banana cultivation presents specificities that make this mapping difficult. Indeed, banana plants only give one bunch by differentiation of their unique terminal bud, and the following cycle of production is insured by selection of one sucker by plant (Tixier *et al*., 2004). Floral induction is not synchronized by seasons (Turner *et al*., 2007), and inter plant variability of growth rate and therefore of cycles length, induces a variability into the phenologic stages present at the within field level at any moment. This variability increases with production cycles and the production becomes approximately continuous. Such a crop is called "asynchronous" by Tixier *et al*. (2004).

Mapping of crop properties, and first of all, yield information, triggered the implementation of precision agriculture to optimize decisions (McBratney *et al*., 2005; Plant, 2001). This mapping cannot be set up as easily for crops that are not season synchronized (Stoorvogel *et al*., 2015) such as banana.

In order to make the harvest easier, banana farmers overcame the issue of variability in the phenological stages by identifying each bunch at bloom by a color tag (Lassoudière, 2012). This system makes it possible to know at any time the age of bunches in the fields with a precision of a week or of half a week. It helps farmers to plan and perform operations, such as deflowering, fruit removal, harvest, sucker selection, according to phenological status.

However, even if technical operations are done plant by plant, decision procedures are defined at farm scale. At a given decision date, the same procedures and crop management decision rules are applied to each plant with the same color tag. Therefore, precision agriculture in banana would aim to optimize farming decision procedures at a lower scale. For example, the decision rules could be adapted plot by plot if pertinent for agronomical reasons. The motivation for precision agriculture is to take into account spatial patterns and specific local conditions. Implementing precision agriculture in banana crops then requires designing spatialized indicators, to quantify their spatial variability and need for local decisions, and at last assessing their usefulness for agronomic support.

A difficulty induced by crop asynchronicity is the number of available plants presenting a particular stage at a given observation date on an area. This number might be too low to derive relevant and precise maps. Irregular distributions of phenological stages in space are also problematic. In order to gather more samples on a given area, two strategies may be set up. The first would be to make measurements on the same phenological stage at different dates. In this case,

† E-mail: j.lamour@fruitiere.fr

variations induced by the different dates of observations should be taken into account. This strategy is called *"sample integration over time"* in the rest of the paper. Another strategy would be to make measurements at the same date but on different phenological stages. In this case, it would be necessary to use a model of plant dynamics to simulate what the plant response would be for the phenological stage of interest.

This work proposes a first approach for the mapping of production properties on an asynchronous culture applied to banana crops. It is based on the strategy "*sample integration over time*". The objective of the paper is to (i) illustrate this methodology for an important banana farming indicator, (ii) assess the spatial variability of this indicator and (iii) provide quality analysis of its relevance at the level of a commercial farm.

Material and methods

Choice of an indicator

In this study, Time from Flower to Maturity (TFM) is investigated. It represents the elapsed time in days between the date of Flowering and the date of operational Maturity of banana bunches in a farm. Indeed, this duration is a common information used to forecast productions and to support the management of harvests (Cottin *et al.*, 1987; Jullien *et al.*, 2008).

In banana, operational maturity is usually assessed according to two criteria (Bugaud and Lassoudière, 2005). The first one is based on the diameter of a specified banana finger in the bunch (Vargas-Calvo, 2012), and the second criterion is related to physiological age. All bunches that reach a maximum age are considered operationally mature and need to be harvested. This maximum age is evaluated according to a sum of degree days (Ganry and Chillet, 2008). Time of flowering is recorded by a particular color tag marked on bunches. So, at any given date, bunches having reached the maximum age correspond to those with a particular color tag. All bunches having this color or having reached a chosen finger diameter need to be harvested. The whole harvest decision procedure is carefully monitored to prevent risks of early ripening during shipping (Jullien *et al.*, 2008).

Differences in TFM throughout the year are mainly due to differences of temperature (Ganry and Meyer, 1975). This factor affects growth rate and also harvest age limit. The hypothesis made in the study is that, for bunches with the same flowering date, spatial differences in TFM exhibit vigor characteristics. Indeed, the higher the growth rate, the shorter the time between flowering and the date at which the finger diameter reaches its pre-determined maximum. For bunches flowering at the same time, differences in TFM can only be explained by differences in finger growth and therefore to different spatial conditions influencing this growth since age limit criteria is the same.

In this paper, the TFM for each week of the year is studied at the farm scale and at the within field scale. Throughout the paper, the set of banana plants whose bunches bloomed at the same week will be called a cohort.

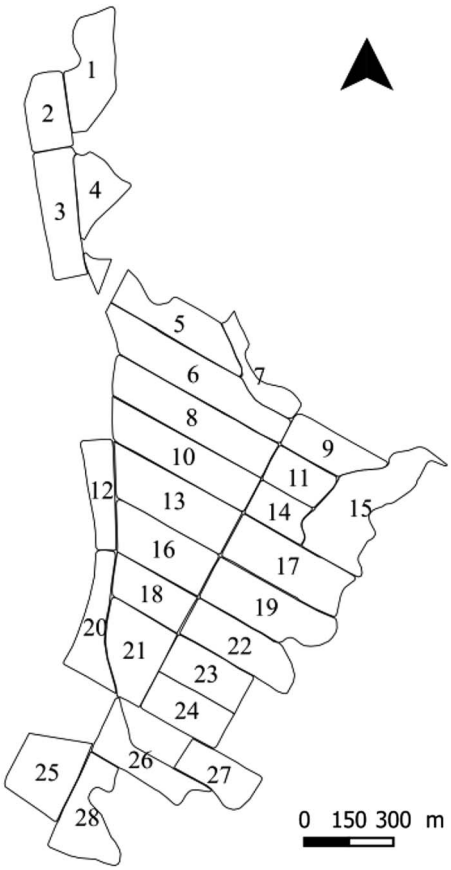

Figure 1 Organization of the farm and its 28 plots.

Study area and data collection

The experiment was performed in a farm in Njombe, Cameroon, during 2015 and 2016.The farm of 92 Ha consisted of 28 contiguous fields (Figure 1). Plots were planted in double rows from years 1999 to 2015. Annual rainfall in 2015 was 3074 mm and average temperature was 26.3°C. Management of the farm (irrigation, fertilization, pest control and fruits care) was done according to the conventional practices in banana farming, as well as the tagging of bunches. Indeed, each inflorescence was manually marked with a color tag at bloom. The marking was organized so that each plot was scouted twice a week. Maturity, based on finger diameter and age limit, was also assessed twice a week. The reference finger for maximum diameter was the central finger of the second hand of each bunch. The diameter limit was the same during all the experimentation.

Operators were equipped with a smartphone that they used to localize bunches when marking blooms and when assessing operational maturity. The spatial accuracy of GNSS system of smartphones was assessed to be 5 meters. Each color tag was identified by a unique bar code and a program was developed to record the time and position during scanning of tags. Finally, each bunch marked with a tag and a bar code was recorded in a database with the following attributes: date of flowering, date of operational maturity, coordinates at flowering and coordinates at operational maturity.

Data analysis and mapping

Data were processed in order to remove global outliers, such as data with wrong coordinates (i.e. outside the field borders) or faulty attribute inputs. The TFM of each bunch was then calculated as the number of days between the date of flowering and the date of operational maturity.

Following the "sample integration over time" strategy, TFM was studied over two consecutive periods P1 and P2, each period lasting 20 weeks and gathering approximately 1500 observations per hectare. P1 started on 13 October 2015 and P2 finished on 16 July 2016. Under such periods of time, the TFM of bunches of different cohorts are likely to be different due to changes in environmental conditions and to farmer's practices that affect all the individuals of a cohort in the farm. Therefore, a measure of *TFM* for an individual *i* may be expressed as:

$$TFM_i = TFM_{cohort(i)} + \varepsilon_i$$

where $TFM_{cohort(i)}$ is the median of all the TFM corresponding to the same cohort and ε_i is the residual. ε_i is a centered TFM and the function ε_i is denominated hereafter "TFMc".

The median was preferred over the mean to estimate the cohort effect because it is less sensitive to the criterion of age limit in maturity assessment. Indeed, if A would be the probability distribution of TFM that would result of a harvest decision based only on the finger diameter, and B be the probability distribution that results from both management criteria, i.e. the finger diameter and the age limit, the mean would not be the same for the two distributions if some TFM values were affected by the age limit criterion. On the contrary, because the age limit criterion can be expected to be greater than the median of A, then the median of both distributions should be the same. The spatial variability of TFMc was quantified both at the whole farm scale and at the within field scale. In both cases all the individuals of each cohort were used to calculate TFM_{cohort}.

The spatial variability of TFMc was quantified with an empirical omni-directional variogram (Bachmaier and Backes, 2008) as described by Bivand *et al.* (2013). To confirm that the increase in TFMc variance with distance was not due to chance, the procedure of Bivand *et al.* (2013) was implemented. It consists in computing variograms after randomly re-assigning measurements to spatial locations. A visually assessment is then made to check if the empirical variogram is out of the range of 95% of the simulated random variograms. All the empirical variograms at the within field scale were analyzed and represented at a maximum lag of 80 meters, corresponding to 80% of the width for most of the plots. This precaution was taken in order to prevent plot shape effects. Otherwise, very elongated plots would possibly induce a directional analysis of variance. The degree of spatial dependence of TFMc was evaluated as the proportion between the spatially correlated variance (C1) and total variance of the variogram (C0 + C1). A variogram model was fitted for both periods P1 and P2 at farm scale. It was only fitted at plot scale for plots and periods where TFMc presented a spatial structure. TFMc were mapped on a square grid with 10 meters side using those theorical variograms by block kriging.

In a second step, the temporal stability of TFMc maps obtained at different periods was studied using a Spearman correlation test. This statistic was preferred over a Pearson test because of the non normality of the data. The P-value is given as an indicator but should not be used here to test statistical hypothesis because of the spatial auto-correlation of the data (Taylor and Bates, 2013).

Tools

Statistical tests and mapping were implemented using R (R Development Core Team, 2008) and packages gstat (Pebesma, 2004) and sp (Pebesma and Bivand, 2005).

Results and discussion

Management of asynchronicity of field data

Figure 2 shows (A) the variability of TFM along cohorts and (B) the effect of centering by median. The total variance of TFM is 39 days² and the intra-cohort variance is 15 days². The number of individuals in a cohort (number of blooms per week) is variable (Figure 3). This number is particularly influenced by flowering peaks of recently planted plots which are still synchronous in the first production cycles. Moreover, as shown in Figure 4, spatial patterns of blooms can be irregular. These peculiarities of banana crops induced by asynchronicity indicate that gathering observations of

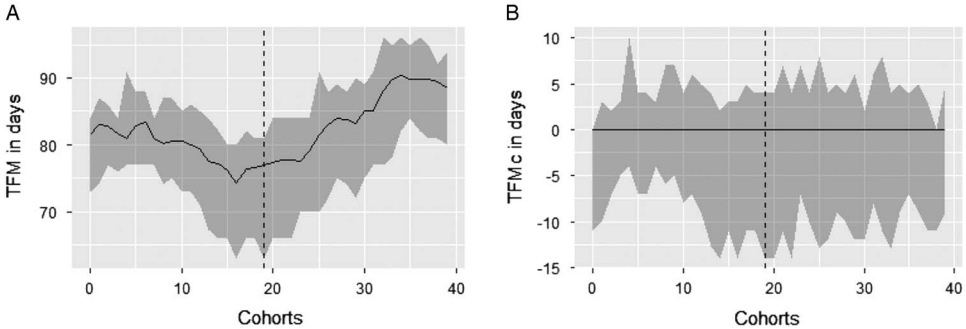

Figure 2 (A) Time between Flowering and Maturity by cohort, (B) centered Time between Flowering and Maturity. Dotted lines correspond to the limits between Period 1 and Period 2. Lines correspond to medians and ribbons contain 95% of the values.

several weeks will improve density and spatial repartition of the observations. They also suggest that TFMc should be analyzed, rather than TFM, in order to compensate variability due to cohorts.

Within farm and within field variability of Time between Flowering and Maturity

Figure 5 and 6 show respectively the variograms corresponding to period 1 and 2 and the maps resulting from the block kriging interpolation. The variograms are out of the range of 95% of the simulated random variograms indicating that the hypothesis of absence of spatial correlation is unlikely. Therefore, the degree of spatial dependence which is 0.30 at period 1 and 0.13 at period 2 is significant. The empirical variogram related to period 2 is not stationary (the variance does not reaches a plateau) which means that the spatial dependence may have a larger scale than 300 meters. Hence, the degree of spatial dependence calculated for this period may be underestimated. Differences in TFMc are visible between plots. For example, the bunches in plot 10 are mature nearly 6 days after the bunches in plots 5 or 6 during period 1. Those results highlight the spatial organization of the TFMc at the farm level, which suggests the relevance to study this indicator for precision agriculture in banana.

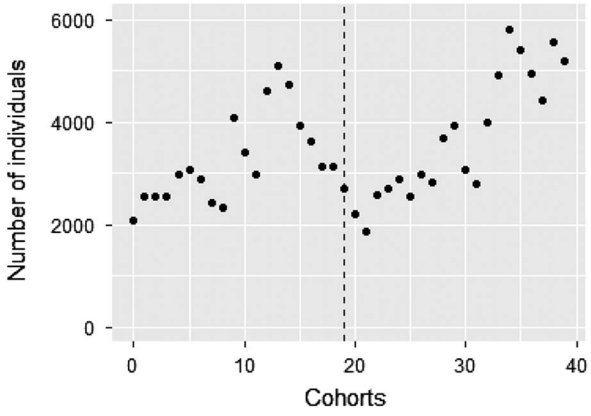

Figure 3 Variation in total number of individuals (= bunches) by cohort on the farm. Each cohort corresponds to a week.

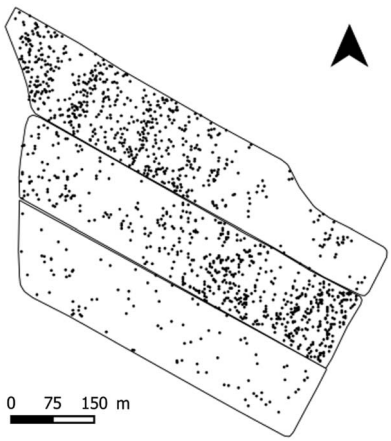

Figure 4 Example of irregular repartition of blooms on plots 6, 8, 10 of week 12.

According to the variogram analyses, twelve plots present a spatial structure for both P1 and P2 periods. Among them, three plots exhibit a temporal stability in their structure (Spearman's Rho >0.5, p-value $<1.10^{-15}$). Thirteen plots exhibit a spatial structure only at period 2. The degree of spatial dependence of TFMc varies from 0% (no structure) to 25%. To strictly respect the hypothesis of stationarity required for kriging procedures, variograms should be modeled at the plot level (Figure 7).

Factors explaining those spatial structures in TFMc might be different according to the situation. Such factors may include biotic (e.g. black sigatoka) and abiotic (light, temperature, soil) causes as well as farm management effects (Srikul and Turner, 1995). More investigations are needed to identify the most influential factors. It is to be noted that those factors may have influenced the development of the fruit directly or indirectly by affecting the plant before bloom. Depending on the season, factors can have different effects on TFMc. The period of sample integration is a compromise between spatial resolution and time resolution of structures that are to be identified. The longer the duration of the analysis, the better the spatial resolution and the smaller the effect of spatial distribution of points, but the lower the opportunity to identify short-live phenomena.

The Nugget effect, i.e., the variance at the origin at the TFMc variogram, is important. It indicates a partial or total absence of correlation between values taken at two very close sites (Arnaud and Xavier, 2000). This can be partially explained by the lack of precision of TFM measurement method. Indeed, observations of flowering and of maturity are done twice a week. This induces a variance of around

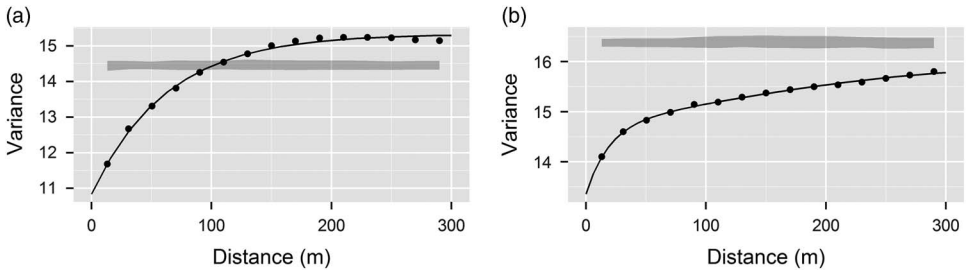

Figure 5 TFMc variograms for period 1(weeks 0 to 19) (a: P1) and period 2 (weeks 20 to 39) (b: P2) at the farm level. Ribbons envelop 95% of the variance of simulated random variograms.

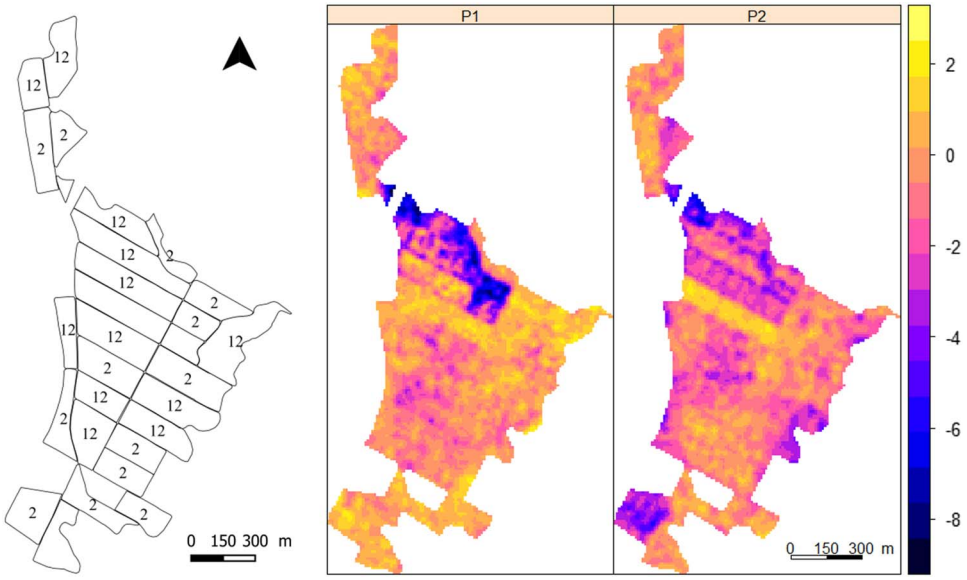

Figure 6 Left: Map of the plots where numbers 1 and 2 mean there is respectively a spatial structure in periods 1 and 2. Right: TFMc maps (in days) obtained by block kriging.

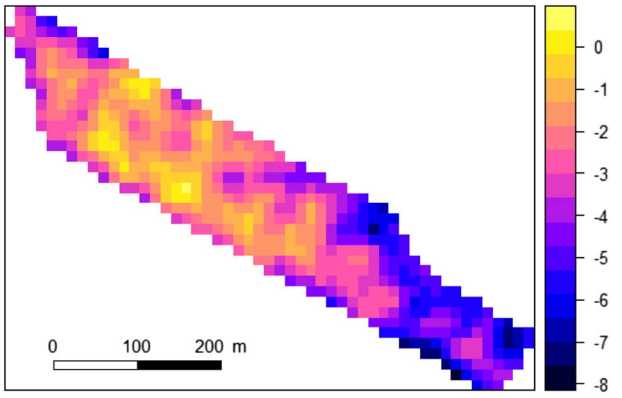

Figure 7 Example of map obtained by kriging at plot level (plot 6).

2.5 days^2 in the measurements (calculation not shown), which represents 17% of the observed variability within a cohort and which is not spatially auto correlated. Moreover, the finger diameter evaluation and the tagging of fruits at bloom are done manually by operators. It can explain a part of the variability in TFM (Bugaud and Lassoudière, 2005). The limited spatial accuracy of the GNSS system, which is 5 meters, might also explain a part of this variability. A high proportion of the Nugget effect also comes from the intra-plant variability which is not negligible.

This strong Nugget effect suggests that both the period of "sampling integration" and block size need to be carefully selected so that the number of observations is sufficient enough to increase the signal to noise ratio when kriging by block (Tisseyre *et al.*, 2015).

Conclusion

Banana production is asynchronous, which is reflected in the heterogeneity of the stages of plant development in the fields. In order to provide map of crop properties with a sufficient spatial resolution, it is necessary to normalize crop data in order to compare observations taken at different times or on different phenological status. This study illustrates the specificities of the asynchronicity that must be taken into account to produce maps of agronomic interest. A method is presented on a maturity indicator that can be interpreted as an index of vigor of fruit growth. The aim of this method is to compensate the temporal variability induced by the gathering of maturity measurement made at different dates and to create a new time-independent indicator. The spatial structure of this indicator was studied. It is spatially structured, on the scale of the farm but also of the small plots. This shows that the functioning of this culture depends on space and that precision farming approaches can be envisaged. In order to improve the agronomic interpretation of these maps so that they can serve as decision support in production, it is still necessary to identify the agronomic factors that explain the observed differences of maturity.

References

Arnaud M and Xavier E 2000. Estimation et interpolation spatiale (Spatial estimation and interpolation). Hermes.

Bachmaier M and Backes M 2008. Variogram or semivariogram? Understanding the variances in a variogram. Precision Agriculture 9, 173–175.

Bivand R, Pebesma E and Gómez-Rubio V 2013. Applied spatial data: analysis with R., 2nd edn, Springer, New York, USA.

Bugaud C and Lassoudière A 2005. Variabilité de la durée de vie verte des bananes en conditions réelles de production (Variability in the green shelf life of bananas in real conditions of production). Fruits 60, 227–236.

Cottin R, Melin P and Ganry J 1987. Modélisation de la production bananière. Influence de quelques paramètres en Martinique. Fruits 42, 691–701.

FAO 2013. FAOSTAT, Production Quantities, Bananas.

Ganry J and Chillet M 2008. Methodology to forecast the harvest date of banana bunches. Fruits 63, 371–373.

Ganry J and Meyer JP 1975. Recherche d'une loi d'action de la temperature sur la croissance des fruits du bananier. Fruits 30, 375–392.

Jullien A, Chillet M and Malezieux E 2008. Pre-harvest growth and development, measured as accumulated degree days, determine the post-harvest green life of banana fruit. The Journal of Horticultural Science and Biotechnology 83, 506–512.

Lassoudière A 2012. Le bananier: Un siècle d'innovations techniques. Editions Quae.

McBratney A, Whelan B, Ancev T and Bouma J 2005. Future directions of precision agriculture. Precision Agriculture 6, 7–23.

Pebesma EJ 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30, 683–691.

Pebesma EJ and Bivand RS 2005. Classes and methods for spatial data in R. R News 5, 9–13.

Plant RE 2001. Site-specific management: the application of information technology to crop production. Computers and Electronics in Agriculture 30, 9–29.

R Development Core Team 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Stoorvogel JJ, Kooistra L and Bouma J 2015. Managing soil variability at different spatial scales as a basis for precision agriculture. Advances in Soil Science pp. 37–72.

Taylor JA and Bates TR 2013. A discussion on the significance associated with Pearson's correlation in precision agriculture studies. Precision Agriculture 14, 558–564.

Tisseyre B, Geraudie V and Saurin N 2015. How to define the size of a sampling unit to map high resolution spatial data?, in: Precision agriculture'15. Wageningen Academic Publishers, pp. 101–113.

Tixier P, Malezieux E and Dorel M 2004. SIMBA-POP: a cohort population model for long-term simulation of banana crop harvest. Ecological Modelling 180, 407–417.

Vargas-Calvo A 2012. Grosor Del Fruto De La Última Y Segunda Mano Como Criterio De Cosecha En Banano (Diameter of the fruit of the last and second hand as criterion of harvest in banana). Agronomía Mesoamericana 23, 41–46.