MEASURING COLOR HUE IN 'SUNRISE SOLO' PAPAYA USING A FLATBED SCANNER[1]

THIAGO DE PAULA OLIVEIRA[2], SILVIO SANDOVAL ZOCCHI[3], ANGELO PEDRO JACOMINO[4]

ABSTRACT- Color hue is a variable frequently used postharvest to assess the physiological maturity of various fruits. In general, mean color hue is quantified by visual techniques, but this method, based on human grading, is tedious and may be erroneous. The color of fruit peel is frequently determined at four points randomly distributed on the equatorial region of a fruit, but this practice can lead to biased results because these points represent the equatorial region only and not the total area of the fruit peel. Consequently, this bias is not due to equipment measurement errors, but to the bias in the sampling process of the points on the fruit epidermis. An alternative to such methods is the digital image which provides information about all regions of the fruit peel, and results in a more accurate mean hue. However, this technique requires calibration to correct the values of luminance, hue and intensity obtained through the scanner. This calibration can be performed from color patterns such that each pattern is assessed by means of a scanner and a colorimeter. Thus, an experiment was conducted using 297 color patterns based on the Munssel color chart for plant tissue. The results showed that the scanner could be used as a device for color assessment where the determination coefficients were above 0.9 for all color components. Next, a second experiment was conducted in order to compare the scanner and colorimeter methodologies. For this, we used a papaya fruit, cv. Sunrise Solo, which was assessed by these two devices for a period of 19 days. The results showed that the image analysis measures have a different mean hue when compared with the mean hue obtained by the method using a colorimeter. Therefore, it is recommended that digital image analysis be used for the evaluation of the hue of fruit peel color when fruit presents non-uniform coloration.

Index terms: Postharvest, hue color, circular data, digital image, colorimeter.

DETERMINAÇÃO DA COLORAÇÃO DE MAMÃO 'SUNRISE SOLO' COM O USO DE SCANNER DE MESA

RESUMO- A coloração é uma variável muito utilizada na pós-colheita para avaliar o estádio de maturação de diversos frutos. Geralmente, a coloração média é quantificada por meio de métodos subjetivos baseados na visão humana, levando a erros graves na classificação dos frutos. Outra forma de quantificar a cor média de um fruto é a partir da amostragem de pontos na superfície de sua casca, processo este geralmente realizado com o auxílio de um colorímetro. Para a análise da coloração do mamão, é usual o pesquisador amostrar quatro pontos na região equatorial da sua casca, porém tal prática pode levar a resultados viesados, pois esses pontos podem representar apenas a região equatorial em vez da região total. Uma alternativa a esses métodos é a utilização de imagens digitais que, no entanto, necessita de calibração para a correção dos valores de luminosidade, tonalidade e intensidade da cor obtidos por meio do *scanner*. Essa calibração pode ser realizada a partir de padrões de cores de tal forma que cada padrão é avaliado por meio do *scanner* e do colorímetro. Assim, foi conduzido o primeiro experimento utilizando 297 padrões de cor baseados na carta de Munssel para tecidos vegetais, o qual mostrou que o scanner pode ser utilizado como dispositivo para a quantificação da cor. Em seguida, foi conduzido um segundo experimento com o objetivo de comparar as metodologias do *scanner* e do colorímetro. Para isso, utilizou-se de um fruto de mamão cv. Sunrise Solo, o qual foi avaliado por meio dos dois equipamentos, por um período de 19 dias. A análise de imagens digitais levou a medidas diferentes de tonalidade média quando comparadas às tonalidades médias obtidas por meio do colorímetro. Sendo assim, é recomendada a utilização de análise de imagens digitais para a avaliação da tonalidade média da coloração da epiderme do mamão 'Sunrise Solo', principalmente nos estádios iniciais e finais de maturação.

Termos para indexação: pós-colheita, tonalidade da cor, dados circulares, imagens Digitais, colorímetro.

[2]Eng. Agr. Doutorando, Curso de Pós-Graduação em Estatística e Experimentação Agronômica, ESALQ/USP. Email: thiago.paula. oliveira@gmail.com. Auxílio FAPESP.
[3]Prof. Doutor, Depto. de Ciências Exatas, ESALQ/USP. Email: sszocchi@usp.br.
[4]Prof. Doutor, Depto. de Produção Vegetal, ESALQ/USP. Email: jacomino@usp.br.

INTRODUCTION

The worldwide purchase of papaya (*Carica papaya* L.) has shown significant growth over the last few years. In 2014, production in Brazil was approximately 1.6 million tonnes of fruit from a harvested area of approximately 32 million hectares (IBGE, 2014).

After the harvest, the quality of the papaya can decline considerably, due to adverse environmental and physical conditions during transportation, storage, distribution and sales. Moreover, extreme or volatile temperatures combined with mechanical damage result in low quality fruits which directly affects their color (FABI et al., 2010).

Peel color is considered the first variable that customers assess and is associated with specific taste or use which may determine its acceptance or rejection in the market (PICHA, 2006). Furthermore, it can be used to determine the maturation stages associated with the harvesting and consumption features of several fruits (FAGUNDES; YAMANISHI, 2001).

In order to standardize the color hue measurement process, simple analysis techniques have been developed, such as digital image analysis, which have overcome the deficiencies of the most commonly used methodologies in post-harvest, namely the ones which use visual scales or a colorimeter (MENDOZA; AGUILERA,2004; DARRIGUES et al., 2008). Moreover, they offer a way to measure other physical attributes of the fruit, such as its shape (RODRÍGUEZ et al., 2010).

In a study involving the assessment of the color of pork, O'Sullivan et al. (2002) concluded that the analysis of digital images can lead to more representative measurements of the color of the meat when compared to a colorimeter, because the images furnish information drawn from the entire surface of the meat.

In 2004, Mendoza and Aguilera proposed the application of digital image analysis to the classification of bananas, when this fruit presents significant natural variability in samples from the same batch. Such variability was also observed by Oliveira et al. (2002) in papaya cv. Golden.

This indicates that there is great potential for using digital images to differentiate the many maturation stages of several fruits, where digital images of all a fruit's peel region can furnish information about this natural variability, and may therefore be taken into account in a statistical model.

It is also important to emphasize the work done by Segnini et al. (1999), who developed a procedure for quantifying color in potatoes using an analysis of images obtained by video; by Yam and Papadakis (2003), whose study showed the advantages, disadvantages and applications of the different artificial vision systems; and by Darrigues et al. (2008), who compared the measurements obtained by different scanners with those of a colorimeter in assessing color patterns, when applying image analysis methodology to the pulp of several fruits.

When using artificial vision systems, such as the scanner, each pixel of the image is represented by combinations of the components *R* (red), *G* (green) and *B* (blue), and each one of them belongs to the [0,1] interval or to the {0,1,2,…,255} set.

To verify the relationship between the *R*, *G* and *B* color components obtained through a colorimeter and a scanner, Segnini et al. (1999) carried out an experiment in which the same color pattern was assessed by both devices and concluded that the relationship is practically linear. The coefficients of determination, R^2, were always greater than 0.92.

However, the use of the *RGB* color space can lead to problems when the intention is to compare object colors as if the comparison had been done by human vision. For example, let (R_1, G_1, B_1) and (R_2, G_2, B_2)) be two colors belonging to the *RGB* color space. The Euclidian distance between them is given by $\sqrt{(R_2-R_1)^2 + (G_2-G_1)^2 + (B_2-B_1)^2}$. Now suppose that two colors are separated by a fixed distance *d*. Depending on the location of the colors in the *RGB* space, the human eye may or may not be capable of distinguishing them.

To correct this problem, the devices used to measure color usually use color systems created by the Commission Internationale de L'Eclairage (CIE) in 1976 (COMMISSION INTERNATIONALE DE L'ECLAIRAGE, 1978). Such systems were created from criteria based on the perception of the human eye. In the *CIEL*a*b* color space, the *L** component indicates the color luminosity, varying from 0 (black) to 100 (white) and the *a** and *b** components belong to the [-60, 60] or [- 100,100] interval. In general, -*a** indicates the color green and +*a* the color red, +*b* the color yellow and +*b** the color blue. In the *CIEL*C*h* system, the color components are given by *L**, *C** and *h**, with *L** being the luminosity (*L** Є [0,100]), *C** the intensity (*C** Є [0,60] or *C** Є [0, 100]) and *h** the hue (*h** Є [0,360º]).

On the other hand, Darrigues et al. (2008) state that one of the problems of digital image analysis is the need to calibrate the equipment (scanner, video cameras, etc.). However, this is usually a procedure that is easy to carry out.

Mendoza and Aguilera (2004), for example,

suggested the use of banana samples at different stages of maturation, in such a way that each banana was assessed by a colorimeter and a camera to obtain the images. Next, through linear regression analysis, the authors obtained the calibration curves for the $L*$, $a*$, and $b*$ components. However, the calibration method presented a low coefficient of determination for the b color component ($R^2 = 0,609$).

To solve this problem, Darrigues et al. (2008) suggested the use of 28 color patterns from Munsell's chart (MUNSELL COLOR COMPANY INC, 1952) to measure the color of tomato. From these color patterns, they obtained determination coefficients greater than 0.92 for the $L*$, $a*$ and $b*$ components.

However, in post-harvest, the color hue is the most cited component in the literature when assessing color in fruit (HUANG et al., 2006; SMILANICK et al., 1995; PULUPOL et al., 1996; PÉK and HELYES, 2010; SILVA et al, 2013) and the analysis of this color component is carried out as if it were a measurement belonging to the real line. According to Zar (2010), this type of analysis can lead to wrong conclusions, since the hue is an angular measurement, described in radians or degrees, and consequently belongs to a circular scale. Thus, analyses of the color hue must be carried out following the specific methodologies applied to circular data, such as those proposed by Mardia (1972), Batschelet (1981) and Fisher (1996).

For example, consider that two red color points were sampled on the same fruit, and that the hues that describe these points are 10° and 350°. Then, the arithmetic mean is 180°, corresponding to the green hue. Using the circular mean, however, yields $X = 1/2$ (sen 10º + sen 350º) = 0 and, $Y = 1/2$ (cos 10 º + cos 350º) = 0 and, consequently, $h = 180/л 0 = 0$º, located in the trigonometric circle (or color circle) exactly between the 10º and 350º values. Therefore, the circular mean should be used as the use of the arithmetic average can provide incorrect results for the mean hue.

In our study, we presented initially the calibration curves for the determination of the peel or pulp color of any fruit based on the color patterns of Munsell's color charts for plant tissue (MUNSELL COLOR COMPANY INC, 1952). Next, we presented descriptive statistics techniques and specific statistical inference procedures for circular data, such as those developed by Batschelet (1981), Fisher (1996) and Jammalamadaka and Sengupta (2001). We then applied such techniques to the peel hue data of papaya (*Carica papaya* L.) cv. Sunrise Solo. Finally, we compared the instrumental methodology (colorimeter) to the digital image analysis methodology (flatbed scanner).

MATERIAL AND METHODS

The experiments were carried out in 2010 at the Department of Exact Sciences and Vegetal Production of ESALQ/USP, Piracicaba, São Paulo. An HP Scanjet G2410 (HP, Palo Alto, CA) flatbed scanner with a resolution of 200 pixels per inch (PPI) and a Minolta CR-300 (KONICA MINOLTA, 1991) colorimeter were used.

Experiment 1: 297 color patterns based on Munsell's color chart of vegetal tissues were used (Figure 1) and each pattern was measured by a colorimeter and a scanner. It is important to mention that this chart was projected to reproduce the exact color of vegetal tissue which is essential to the calibration procedure.

Two points were observed at the central region of each color pattern using the colorimeter. Thus, for each pattern the arithmetic mean of the $L*$, $a*$ and $b*$ and $C*$ components were calculated based on the respective observed values of each one of them. Next, the mean value of the $h*$ component was calculated using the circular mean of observations.

Then, each color pattern was digitalized by the scanner and saved as a *TIFF* image file (.tif), which, despite having a larger file size than the jpg format, preserves the original quality of the image and is recommended when working with dark objects (RODRÍGUEZ et al., 2010).

Experiment 2: One papaya fruit from the cv. Sunrise Solo was stored in a cooling chamber at 18°C at 80% ± 5% relative humidity. The fruit's peel was assessed daily over a period of 19 days using the scanner at a resolution of 200 PPI and the colorimeter. Because the interest of this experiment lies in the comparison between the scanner and colorimeter methodologies when assessing the mean hue of the fruit peel, the use of only one papaya fruit is sufficient, given that both devices were used to assess the same fruit.

Throughout the assessments made with the colorimeter, four points in the equatorial region (Figure 4) were sampled, as recommended by Oliveira et al. (2002).

When assessing the fruits using the scanner, both faces of the fruit were digitalized, one being named face A and the other, face B (Figure 2). Each image was saved as a TIFF file. In general, the image of the whole peel region provided around 200 to 250 thousand pixels.

In both experiments, to avoid external light interference when the object was digitalized, a cardboard box coated internally and externally with

black cloth was built, as recommended by Darrigues et al. (2008).

The use of the box also minimized shadowing effects, as well as providing a dark background for the image. We were careful to ensure that color luminosity, intensity and hue did not coincide with those of the object of interest. In practice, the use of a black cloth was proposed because as the luminosity is closer to zero, it is easier to separate the object from the dark background.

The selection of the object of interest can be made using the autoThreshold function from the rtiff package version 1.4.1 (KORT, 2012) of the R computational environment (R CORE TEAM, 2016). Hence, for each pattern K pixels from the digital images were considered, from which the R_k, G_k and B_k values for each pixel were read ($k = 1,2,3, ..., K$)

After that, the conversion from the RGB color space to $CIEL*a*b*$ and from to $CIEL*a*b*$ $CIEL*C*h*$ had to be carried out. It is important to note that the method of conversion from RGB values to $CIEL*a*b*$ depends on the type of illumination, angle of observation and type of RGB color space. In this case, we used $D65$ illumination with an observation angle equal to 10° and standard RGB color space ($sRGB$). This conversion must be performed, according to Darrigues et al. (2008) and Zar (2010) in three steps:

i. Conversion from values of the RGB system to the XYZ system.

Let (R_k, G_k, B_k) be the color of the k-th pixel of the image, $k = 1,2,...K$. We then have

$$\begin{bmatrix} X_k \\ Y_k \\ Z_k \end{bmatrix} = M_\tau \begin{bmatrix} f(R_k) \\ f(G_k) \\ f(B_k) \end{bmatrix}$$, where M_t is given by:

$$M_\tau = \begin{bmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721750 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{bmatrix}$$ and

$f(.)$ is a function given by :

$$f(x_k) = 100 \left(\frac{\frac{x_k}{255} + 0.055}{1.055} \right)^{2.4} \text{ if } x_k \in \{0,1,2,\ldots,255\}$$

or

$$f(x_k) = 100 \left(\frac{x_k + 0.055}{1.055} \right)^{2.4} \text{ if } x_k \in [0,1]$$

ii. Conversion from values of the XYZ system to $CIEL*a*b*$.

In this case,

$$L_k^* = 116\, g\left(\frac{Y_k}{Y_n}\right) - 16 \quad a_k^* = 500 \left[g\left(\frac{X_k}{X_n}\right) - g\left(\frac{Y_k}{Y_n}\right) \right]$$

and , $b_k^* = 200 \left[g\left(\frac{Y_k}{Y_n}\right) - g\left(\frac{Z_k}{Z_n}\right) \right]$, with

$$g(t_k) = \begin{cases} \frac{1}{3}\left(\frac{29}{6}\right)^2 t_k + \frac{4}{29}, & \text{if } t_k \leq \left(\frac{6}{29}\right)^3 \\ 1 t_k^{\frac{1}{3}}, & \text{if } t_k > \left(\frac{6}{29}\right)^3 \end{cases}$$

where Y_n, X_n and Z_n are values that depend on the type and illumination angle of the scanner. Hence, for the $D65$ illumination and a 10° angle, used in this work, $X_n = 95,047$, $Y_n = 100.000$ and $Zn = 108.883$. Note that the values of $a*$ and $b*$ belong to the interval from -100 to 100 and, then, to transform to the interval from -60 to 60, we must multiply the values of these components by 60/100.

iii. Conversion from the values of the $CIEL*a*b*$ system to values of the $CIEL*C*h*$ system:

and $C_k^* = \sqrt{a_k^{*2} + b_k^{*2}}$

$$h_k^* = \frac{180}{\pi} \times \begin{cases} 0, & \text{if } a_k = 0 \text{ and } b_k = 0 \\ \frac{\pi}{2}, & \text{if } a_k = 0 \text{ and } b_k = 1 \\ \frac{3\pi}{2}, & \text{if } a_k = 0 \text{ and } b_k = -1 \;\; (1) \\ \text{arctg}\frac{b_k}{a_k}, & \text{if } a_k > 0 \text{ and } b_k > 0 \\ \pi - \text{arctg}\left|\frac{b_k}{a_k}\right|, & \text{if } a_k < 0 \text{ and } b_k > 0 \\ \pi + \text{arctg}\frac{b_k}{a_k}, & \text{if } a_k < 0 \text{ and } b_k < 0 \\ 2\pi - \text{arctg}\left|\frac{b_k}{a_k}\right|, & \text{if } a_k > 0 \text{ and } b_k < 0 \end{cases}$$

where h_k* is presented in degrees and $k = 1,2 ..., K$.

Once the (L_k^*, a_k^*, b_k^*) or (L_k^*, C_k^*, h_k^*) colors are obtained for each one of the K pixels, we must calculate the average values for each of their components. For components $L*$, $a*$, $b*$ and $C*$, Zar (2010) suggests the use of the arithmetic mean and for component $h*$, the circular mean ($\bar{h}$) , that can be obtained by replacing a_k^* by $\overline{X} = 1/k \sum_{k=1}^{K} \sin(\bar{h}_k^*)$ and b_k^* by $\overline{Y} = 1/k \sum_{k=1}^{K} \cos(h_k^*)$ in expression (1).

Given a specific color component the i-th pattern of the color chart $(i = 1,2...,M)$ and its average related to the scanner and the colorimeter were now considered, and a table such as Table 1 was formulated.

Once the M mean values had been obtained by the scanner and the colorimeter for each color component, we proceeded to the final step of calibration, in which the calibration functions were obtained through linear regression.

Let h_1, h_2 be a set of $N \leq K$ data from the random variable h^*. We will present now how to summarize the data using frequency distributions and a sector histogram.

The sector histogram or rose diagram is the graphical representation of the j-th frequency, $j = 1,2,...,T$, given by the angle sector a $(j-1)$ to ja and length $\sqrt{f_j}$, where $a = 360/T$ is the amplitude of the T intervals, as shown in Table 2.

Note that in this case the area of this sector is given by $af_j/2$, i.e., proportional to f_j. These plots can be constructed through functions of the libraries "CircStats" and "circular" (JAMMALAMADAKA; SENGRUPTA, 2001), as implemented by the R software program (R CORE TEAM, 2016).

With regard to location measurements, the circular mean, $\overline{h}$, has been described previously. As a measurement of angular concentration, we may consider the length (r) of the vector ($\overline{OP}$) in which O has coordinates ($0,0$) and P has coordinates (X,Y).

This measurement can be obtained through the expression $\sqrt{X^2 + Y^2}$ whose value belongs to the [0,1] interval, equal to 1 when all the circular data are equal, or 0 when all of them are distributed with minimum concentration. It is important to note that both $\overline{h}$ and r are frequently presented graphically by the already-mentioned OP vector.

Once r is a measurement of angular concentration, we may consider, according to Zar (2010), $1 - r = 0$ as a measurement of dispersion for circular data, in which $1 - r = 0$ shows zero dispersion and $1 - r = 0$ shows maximum dispersion.

This measurement was defined by Mardia (1972) as the circular variance S^2 given by $S^2 = 1 - r$. Batschelet (1981), on the other hand, defined the angular variance dispersion measurement by $S^2 = 2$ $(1 - r)$ with values belonging to the [0,2] interval.

Note that the dispersion measurements presented so far have a defined maximum limit. Hence, for the new statistic to have values belonging to the $[0, \infty]$ interval, such as in the linear case, Mardia (1972) proposed the variance measurement $s^2_0 = -2\ln r$, for $r \in \,]0,1]$.

However, in such cases where the dispersion is the largest possible, $S^2 = 2$, or $s^2_0 \to \infty$, the distribution of the observed points is not necessarily uniform across the circumference. It is important to observe that the averages S, s and s_0 are expressed in radians, and to express those in degrees, all one needs to do is to multiply them by $180/_\Pi$.

Let Θ be the populational angle, $(1-\alpha)100\%$ the confidence level and $\overline{h}$ the sample circular mean. Therefore, the confidence interval for Θ is given by $IC(\Theta)_{(1-\alpha)100\%} = \overline{h} \pm d$, where d, according to Zar (2010), and for small samples such as , as is usual in assessments of hue with the colorimeter, is given by

$$d = \arccos\left[\frac{1}{Nr}\sqrt{\frac{2N(2R^2 - N\chi^2_{\alpha,1})}{4N - \chi^2_{\alpha,1}}}\right],$$

if $\sqrt{\frac{\chi^2_{\alpha,1}}{2N}} < r \leq 0,9$ or

$$d = \arccos\left[\frac{1}{Nr}\sqrt{N^2 - (N^2 - R^2)\exp\left(\frac{\chi^2_{\alpha,1}}{N}\right)}\right],$$

if , in which N is the sample size and $X^2_{\alpha,1}$ is the quantile of order $(1-\alpha)100\%$ of the chi-squared distribution with one degree of freedom.

RESULTS AND DISCUSSION

From the first experiment, the calibration functions for each color component (L^*, a^*, b^*, C^*, and h^*), as well as the respective coefficients of determination, were used as a measurement of model goodness-of-fit. The coefficients of determination were greater than 0.9 for all color components (Figures 3 and 4), i.e., generally the scanner may be used as a device to assess these components since the range of pulp or peel colors can be represented by Munsell's color chart for vegetal tissues. This result agrees with the results presented by Darrigues et al. (2008), once the proportion of variation of the response variable that can be explained by the regression was considered high ($R^2 \geq 0,9$).

In Figure 5, an arrow was added to the rose diagrams extending from the center of the circumference, indicating the size and direction of the mean vector and generic points (in red) of the analyzed fruit on days 1, 10 and 19 using the scanner and the colorimeter. From this Figure and Table 3 it can also be seen that the initial mean hue of the fruits assessed by the colorimeter and the scanner were similar, with $\overline{h} \cong 112,96°$ (green) for the scanner and $\overline{h} \cong 115,41°$ (green) for the colorimeter.

On the other hand, the final mean hue was apparently different between the methodologies, varying between a yellow fruit color presented by the scanner ($\overline{h} \cong 91,37°$) and an orange fruit color presented by the colorimeter, ($\overline{h} \cong 75,52°$).

Such difference can be due to both the fact that the colorimeter is a punctual device and the sample region, as well as the number of observed points. It is important to note that the fruit presented a gradual maturation over time independent of the equipment as given by the circular means (Table 3).

Furthermore, there is a high angular concentration of the points ($r \cong 1$)), i.e., the dispersion of the points in the circumference is small for both methodologies, implying that the hue variability is small for both methodologies. This is since the papaya naturally presents hues varying in a restricted range of hues, generally between [45º, and 150º] , and also because the hue changes gradually over time.

Additionally, it is important to note that in image analysis it is common to find a small percentage of pixels with hues outside of the green to yellow-orange interval, going clockwise. Such points may occur due to small stains in the fruits, which can be caused by disease or mechanical damage.

Since there may be points away from the region of major data concentration, they represent only 0.8% Thus, they practically do not influence the mean hue calculated by the circular mean. However, if the mean hue were to be calculated by the arithmetic mean, this result would be wrong because there are points with the hue close to 0º. Moreover, the analysis of digitalized images makes it possible to obtain a mean hue closer to the true mean, the mean hue resulting from the colorimeter methodology. This is related to sample size, as well as the position of each point observed in the fruit peel region.

Another important physiological feature observed in this experiment was that the fruit ripened from yellow bands in the peel originating from the ESTILAR region towards the PEDUNCULAR insertion, according to Oliveira et al. (2002). This may lead, apparently, to different variances between the methodologies, mainly at the beginning of the experiment, because, initially, the papaya color is more uniform in the equatorial region of the fruit. In the ESTILAR region of the fruit, the color varies due more to the formation of yellow bands in its peel.

This way, when analyzing the angular deviation, circular deviation and angular standard deviation, presented in Table 3, it was clear that they differ between the methodologies at the beginning of the experiment, and generally they are greater when the scanner is used compared to the colorimeter. This shows that when assessing the fruit in the equatorial region only it is not possible to capture the natural variability of the hue in the fruit peel. Therefore, the analysis of digital images emerges as an important tool when the fruit presents non-uniform color in the ripening process.

This can be verified when looking at the confidence intervals for the colorimeter data (Figure 6 and Table 3). Note, for example, there are moments when the fruit is assessed as green by the colorimeter when it is actually at a more advanced ripening stage (green with yellow stains in the ESTILAR region) when assessed by the scanner. In another case, it is assessed as ripened by the colorimeter (yellow-orange, $\overline{h}* = 75$) when it is actually in a less advanced ripening stage (yellow with few orange stains, $\overline{h}* = 85$) when assessed by the scanner. Hence, it is important to build the confidence intervals for the mean hue at the j-th time unit to assess if the measurement error between the methodologies is "big".

However, the construction of these intervals was only necessary for the colorimeter methodology, once the scanner data had provided information on the entire region of the fruit peel. In general, each digital image generated around 200 to 250 thousand observations, and, thus, we can say that as $N \rightarrow \infty$, the variability of the mean hue tends to zero. In other words, the use of images of both faces of the fruit provides information on the populational mean hue.

For each methodology, different regions of the fruit are used to observe the points and this must be considered, because then, in theory, the methodologies cannot be compared. However, in practice, it is common to use the points observed in the equatorial region of the peel to describe the mean hue of the entire surface, implying a probable measurement error.

Thus, lower and upper limits for the 95% confidence interval for the mean hue resulting from the colorimeter methodology were constructed (Table 4). When these intervals were compared to the mean peel hue of the fruit resulting from the scanner methodology, we noted that the methodologies of color analysis presented different results in days 1 to 5 and days 14 to 18, confirming the initial hypothesis that the methodologies can produce different results (Figure 6).

Therefore, the assessment of the mean peel hue using digitalized image analysis techniques is an alternative for standardizing the process of fruit color determination, as suggested by Chen et al. (2002). Whatsmore, the use of statistical techniques for the analysis of circular data allows for the assessment of data concentration, as well as the circular mean and dispersion considering the natural periodicity of the circumference.

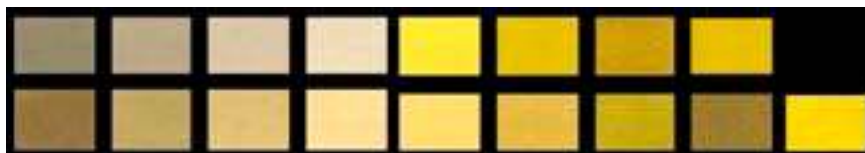As regards future work, we intend to use regression models for longitudinal circular data.

FIGURE 1 – Example of 17 patterns obtained from the Munsell's color chart for vegetal tissue.

FIGURE 2 – Digitalized images of both faces, A and B, of the same papaya cv. Sunrise Solo fruit, as well as the location of the points in the equatorial region of the fruit that were observed using the colorimeter.
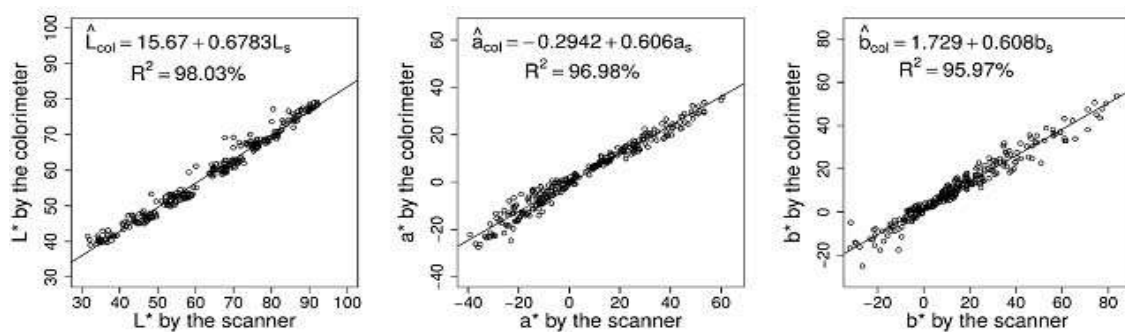

FIGURE 3 – Calibration plots for the scanner methodology for the color components L^* (a), a^* (b) and b^* (c) from experiment 1.
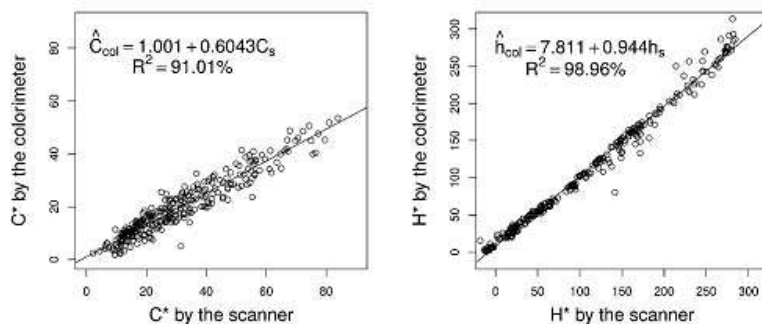

FIGURE 4 – Calibration plots for the scanner methodology for the color components C^* (a) and h^* (b) from experiment 1.

TABLE 1 – Mean values of the color components for each Munsell's color chart standard assessed by the scanner and the colorimeter.

Method	Standard				
	1	2	3	...	m
Scanner	x_1	x_2	x_3	...	x_m
Colorimeter	y_1	y_2	y_3	...	y_3

TABLE 2 – Frequency distribution and length of sectors for the construction of the histograms or rose diagrams.

j	Interval	f_j	$c_j = \sqrt{f_j}$
1	$[0,h^*[$	f_1	$c_1 = \sqrt{f_1}$
2	$[h^*, 2h^*[$	f_2	$c_2 = \sqrt{f_2}$
3	$[2h^*, 3h^*[$	f_3	$c_3 = \sqrt{f_3}$
...	...	...	...
u	$[uh^* - h^*, uh^*[$	f_u	$c_u = \sqrt{f_u}$
...	...	...	...
k	$[360-h^*, 360[$	f_k	$c_k = \sqrt{f_k}$
Total		n	

TABLE 3 – Circular mean ($\overline{h}$), length of the mean vector (r), circular deviation (S), angular deviation (s) and standard circular deviation (s_0) of the peel hue measured by the scanner and the colorimeter over 19 days and lower limits (LI) and upper limits (LS) of the confidence interval for the mean hue measured by the colorimeter.

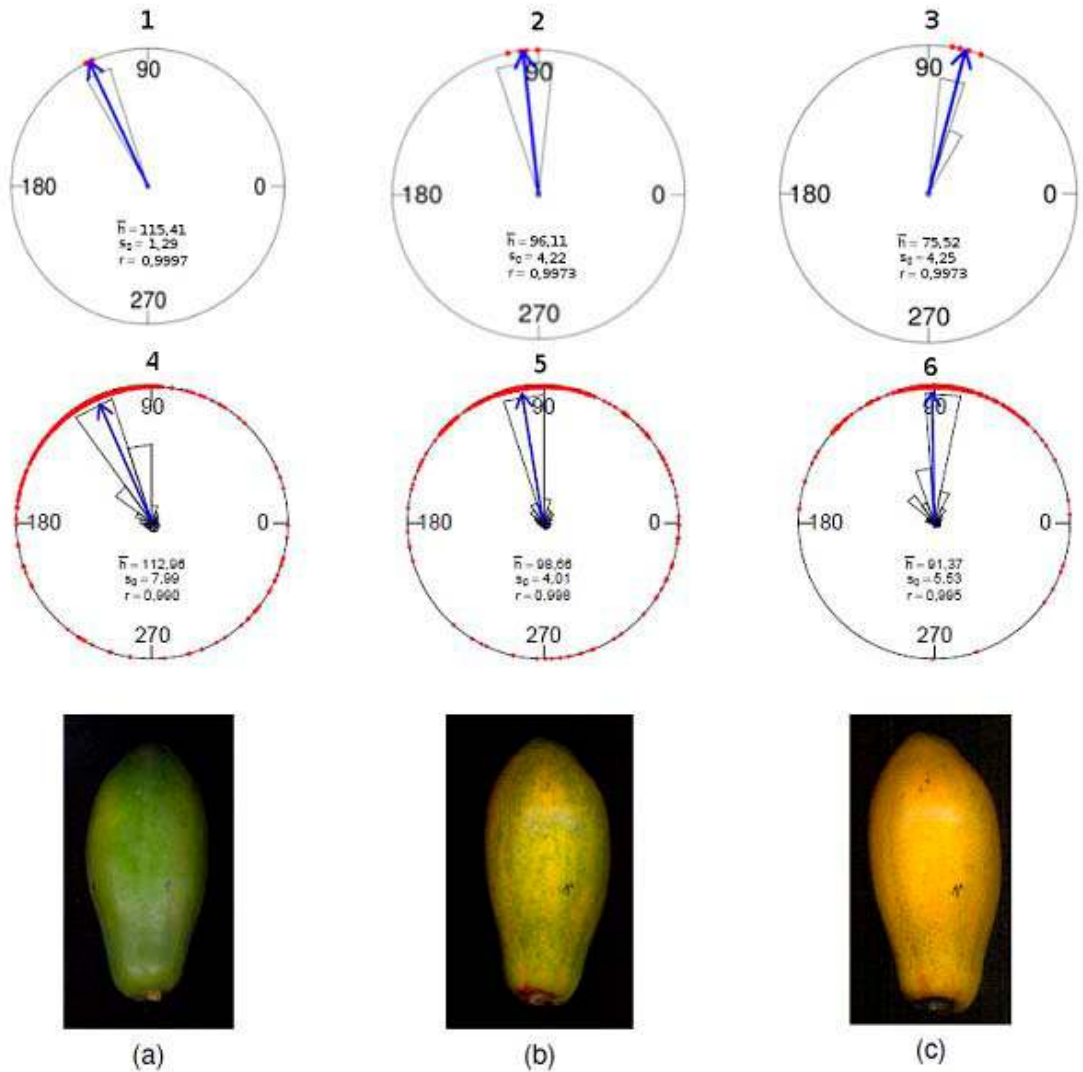| Day	Scanner					Colorimeter						
	$\overline{h}$	S	s	s_0	r	$\overline{h}$	LI	LS	S	s	s_0	r
1	112,96	5,64	7,97	7,99	0,9903	115,41	113,77	117,04	0,91	1,29	1,29	0,9997
2	108,54	3,40	4,80	4,81	0,9965	114,57	112,65	116,49	1,07	1,51	1,51	0,9997
3	106,94	3,28	4,63	4,64	0,9967	112,90	109,58	116,22	1,85	2,61	2,61	0,9990
4	105,23	2,89	4,08	4,09	0,9975	110,95	107,31	114,59	2,02	2,86	2,86	0,9988
5	103,64	5,12	7,25	7,26	0,9920	109,54	105,14	113,93	2,44	3,45	3,46	0,9982
6	102,99	5,47	7,73	7,75	0,9909	106,82	103,50	110,15	1,85	2,61	2,62	0,9990
7	101,01	4,78	6,76	6,77	0,9930	103,41	99,15	107,66	2,36	3,34	3,34	0,9983
8	100,11	2,54	3,59	3,60	0,9980	101,46	96,92	106,00	2,52	3,57	3,57	0,9981
9	98,89	2,39	3,38	3,38	0,9983	98,66	93,97	103,36	2,61	3,69	3,69	0,9979
10	98,66	2,84	4,01	4,01	0,9976	96,11	90,74	101,49	2,98	4,22	4,22	0,9973
11	97,21	2,36	3,34	3,34	0,9983	93,32	86,30	100,34	3,89	5,50	5,50	0,9954
12	96,14	2,36	3,33	3,34	0,9983	90,86	84,76	96,96	3,38	4,78	4,78	0,9965
13	95,15	2,31	3,26	3,27	0,9984	88,05	80,82	95,28	4,00	5,66	5,66	0,9951
14	94,27	4,32	6,11	6,12	0,9943	85,88	78,72	93,03	3,96	5,60	5,60	0,9952
15	93,44	2,30	3,25	3,25	0,9984	83,96	76,78	91,13	3,97	5,62	5,62	0,9952
16	92,85	4,74	6,70	6,71	0,9932	80,92	74,42	87,43	3,60	5,10	5,10	0,9960
17	91,51	4,93	6,97	6,98	0,9926	77,89	72,47	83,31	3,01	4,25	4,25	0,9972
18	91,71	2,90	4,10	4,11	0,9974	76,95	71,74	82,16	2,89	4,10	4,09	0,9975
19	91,37	3,91	5,52	5,53	0,9954	75,52	70,11	80,93	3,00	4,25	4,25	0,9973

FIGURE 5 – Rose diagram adding a vector of length *r* with direction given by $\overline{h}$ and representation of the hue in each point (in red), h_i^* (i = 1,2,3,4), of the analyzed fruit in days 1 (a), 10 (b) and 19 (c) using the colorimeter, plots 1 to 3, and the *scanner*, plots 4 to 6, as well as the digitalized image of one of the fruit's faces in the respective days.
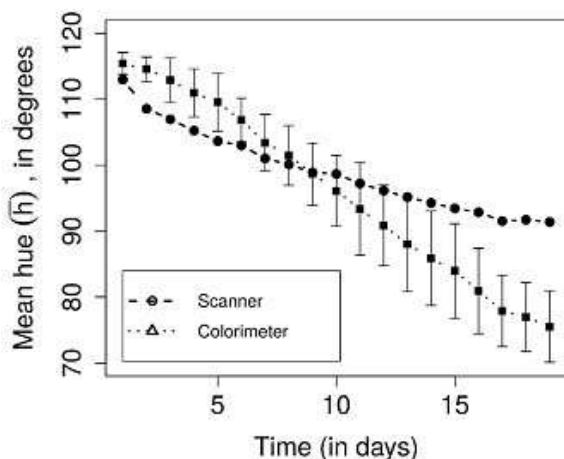
FIGURE 6 – Mean hue of the peel color assessed using the scanner and the colorimeter in a period of 19 days.

CONCLUSIONS

The scanner can be used as a device in the determination of the components L^*, a^*, b^*, C^* and h^* of a fruit peel or pulp color belonging to the Munsell's color chart's established intervals.

The instrumental method, using the colorimeter, and the method of digital image analysis, using a scanner, lead to different peel color measurements at the beginning and at the end of the ripening process of papaya.

Therefore, based on the sample size, we conclude that the use of digital images, from devices such as a scanner, to assess the color fruit peel which present non-uniform color, is important because the use of these images makes it possible to work with large samples ($N > 200.000$), which brings the mean hues closer to the true hue.

ACKNOWLEDGEMENTS

REFERENCES

BATSCHELET, E. **Circular statistics in biology**. New York: Academic Press, 1981. 371 p.

CHEN, Y. R.; CHAO, K.; KIM, M. S. Machine vision technology for agricultural applications. **Computers and Electronics in Agriculture**, Beltsville, v.36, p.173–191, 2002.

COMMISSION INTERNATIONALE DE L'ÉCLAIRAGE. **Recommendations on uniform color spaces**: color-diference equations, psychometric color terms. Paris, 1978. 1v.

DARRIGUES, A.; HALL, J.; KNAAP, E van der; FRANCIS, D. M. Tomato analyzer-color test: a new tool for efficient digital phenotyping. **Journal of the American Society for Horticultural Science**, Alexandria, v.133, n.4, p.579–86, 2008**.**

FABI, J.P.; PERONI, F.H.G.; GOMES, M.L.P.A. Papaya, mango and guava fruit metabolism during ripening: Postharvest changes affecting tropical fruit nutritional content and quality. **Fresh Produce**, Isleworth, v.4, n.1, p.56-64, 2010.

FAGUNDES, G. R.; YAMANISHI, O. K. Características físicas e químicas de frutos de mamoeiro do grupo 'solo' comercializados em 4 estabelecimentos de Brasília-DF. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 23, n. 3, p. 541-545, 2001.

FISHER, N.I. **Statistical analysis of circular data**. Cambridge: Cambridge University Press, 1996. 277 p.

HUANG, S.; HART, H.; LEE, H.; WICKER, L. Enzymatic and color changes during post-harvest storage of lychee fruit. **Journal of Food Science**, Champaign, v.55, n.6, p.1762-3, 1990.

HUANG, C.; YU, B.; YUANWEN, T.; SU, J.; SHU, Q.; CHENG, Z.; ZENG L. Effects of fruit bagging on coloring and related physiology, and qualities of red Chinese sand pears during fruit maturation. **Scientia Horticulturae**, v. 121, n. 2, p. 149-158, 2006.

IHAKA, R.; MURRELL, P.; HORNIK, K.; FISHER, J. C.; ZEILEIS, A. **Color space manipulation**. R package version 1.2-4. 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA-IBGE. **Produção agrícola mundial**: culturas temporárias e permanentes. Produção Agrícola Municipal, Rio de Janeiro, v. 41, p. 1-100, 2014.

JAMMALAMADAKA, S. R.; SENGUPTA, A. **Topics in circular statistics**. Chennai: World Scientific Publishing, 2001. 322 p.

KONICA MINOLTA. **Chroma meter CR-300/CR-310/CR-321/CR-331/CR-331C**. Ramsey: Minolta, 1991.

KORT, E. **Rtiff**: a tiff reader for R. Versão 1.4.1. 2012. Disponível em: <https://cran.r-project.org/web/packages/rtiff/rtiff.pdf>.

MARDIA, K.V. **Statistics of data directional**. London: Academic Press, 1972. 377 p.

MENDOZA, F.; AGUILERA, J.M. Application of image analysis for classification of ripening bananas. **Journal of Food Science**, Champaign, v.69, n.9, p. E471–E477, 2004.

MOITA, A. W. **Análise de variância para dados de tonalidade de cor: um caso de dados circulares**. 2004. 50 f. Dissertação (Mestrado em Estatística e Experimentação Agronômica) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2004.

MUNSELL COLOR COMPANY. **Munsell color charts for plant tissues**. Grand Rapids, MI: Munsell Color, 1952.

O'SULLIVAN, M.G.; BYRNE, D.V.; MARTENS, H.; GIDSKEHAUG, L.H., ANDERSEN, H.J.MARTENS, M. Evaluation of pork colour: prediction of visual sensory quality of meat from instrumental and computer vision methods of colour analysis. **Meat Science**, Amsterdam, v.65, p.909-18, 2003.

OLIVEIRA, M.A.B.de; VIANNI, R.; SOUZA, G.de; ARA UJO, T.M.R. Caracterização do estádio de maturação do papaia "golden" em função da cor. **Revista Brasileira de Fruticultura**, Jaboticabal, v.24, n.2, p.559-61, 2002.

PÉK, Z.; HELYES, L. Color changes and antioxidant content of vine and postharvest ripened tomato fruits. **Horticultural Science**, Prague, v.45, n.3, p.466-8, 2010.

PICHA, D. Horticultural crop quality characteristics important in international trade. **Acta Horticulturae**, The Hague, v.712, n.1, p.423–6, 2006.

PULUPOL, L.U.; BEHBOUDIAN, M.H.; FISHER, K.J. Growth, yield, and postharvest attributes of glasshouse tomatoes produced under deficit irrigation. **Horticultural Science**, St Joseph, v.31, n.6, p.926-9, 1996.

R Development Core Team. **R:** A Language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2014.

RODRÍGUEZ, G.; STRECKER, J.; BREWER M.; GONZALO, M. J.; ANDERSON, C.; LANG, L.; SULLIVAN, D.; WAGNER, E.; STRECKER, B.; DRUSHAL, R.; DUJMOVIC, N.; FUJIMURO, K.; JACK, A.; NJANJI, I.; THOMAS, J.; GRAY, S.; KNAAP, E. van der. **Tomato analyzer user manual version 3**. 2010. 26 p. Disponível em: <http://vanderknaaplab.uga.edu/files/Tomato_Analyzer_3.0_Manual.pdf>

SEGNINI, S.; DEJMEK, P. A low cost video technique for colour measurement of potato chips**. Lebensmittel-Wissenschaft & Technologie**, London, v.32, p.216-22, 1990.

SILVA, J. M. da; MIZOBUTSI, G. P.; MIZOBUTSI, E. H.; CORDEIRO, M. H. M.; FERNANDES, M. B. Conservação pós-colheita de pinha com uso de 1-metilciclopropeno. **Revista Brasileira de Fruticultura**, Jaboticabal, v.35, n.4, p.1201-8, 2013.

SMILANICK, J. L.; MARGOSAN, D. A.; HENSON, D. J. Evaluation of heated solutions of sulfur dioxide, ethanol, and hydrogen peroxide to control postharvest green mold of lemons. **The American Phytopathological Society**, St Paul, v.79, n.7, p.742-7, 1995.

YAM, K.L.; PAPADAKIS, S.E. A simple digital imaging method for measuring and analyzing color of food surfaces. **Journal of Food Engineering**, Amsterdam, v. 61, p. 137-142, 2004.

ZAR, J. **Biostatistical Analysis**. New Jersey: Prentice Hall, 2010. 944 p.