

CIE L*a*b* polygon for quantifying scab (*Sphaceloma perseae*) in avocado fruit images

Polígono CIE L*a*b* para cuantificar roña (*Sphaceloma perseae*) en imágenes de frutos de aguacate

Erik Acuayte-Valdes¹; Alejandro F. Barrientos-Priego²; Elisa Duran-Peralta¹; Mercedes Cabrera-Morales³; Graciano Javier Aguado-Rodríguez^{3*}

¹Colegio de Postgraduados. Carretera México-Texcoco km 36.5, Montecillo, Estado de México, C. P. 56264, MÉXICO.

²Universidad Autónoma Chapingo. Carretera México-Texcoco km 38.5, Chapingo, Texcoco, Estado de México, C. P. 56230, MÉXICO.

³Universidad Politécnica de Francisco I. Madero. Francisco I. Madero, Tepatepec, Hidalgo, C. P. 42650, MÉXICO.

*Corresponding author: gaguadoro@conahcvt.mx

Abstract

Avocado scab causes great economic losses in Mexico and the world. This study aimed to provide an automatic method of quantifying this symptom. The proposed method uses polygonal regions from the a-b plane of the CIE L*a*b* color format to identify a color (even with variations) and different avocado fruit shades. The polygons were used to determine the scab percentage in 70 Hass avocado fruit images. The scab percentage was estimated with a diagrammatic scale, and the regression equation $y = 0.5848x$, with an $r^2 = 0.80$, was found when comparing these results with scab percentages obtained from digitized images (observed). With the method presented in this study, the equation $y = 0.9923x$ was found, with an $r^2 = 0.8672$, so the proposed automatic method performed better.

Keywords: *Persea americana*, colorimetric parameters, digital image processing.

Resumen

La roña del aguacate causa grandes pérdidas económicas en México y en el mundo. El objetivo de este estudio fue proporcionar un método automático de cuantificación de este síntoma. El método propuesto utiliza regiones poligonales, del plano a-b, del formato de color CIE L*a*b* para identificar un color (aún con variaciones) y diferentes tonalidades del fruto de aguacate. Con los polígonos se determinó el porcentaje de roña en 70 imágenes de frutos cv. Hass. El porcentaje de roña se estimó con una escala diagramática y se encontró la ecuación de regresión $y = 0.5848x$, con una $r^2 = 0.80$, al comparar estos resultados con porcentajes de roña obtenidos de imágenes digitalizadas (observados). Con el método presentado en este estudio, se encontró la ecuación $y = 0.9923x$, con una $r^2 = 0.8672$, por lo que el método automático propuesto tuvo un mejor desempeño.

Palabras clave: *Persea americana*, parámetros colorimétricos, procesamiento digital de imágenes.



Introduction

Fruit scab is caused by the fungus *Sphaceloma perseae*, and is one of the main phytosanitary problems of avocado (*Persea americana* Mill.), which reduces its export potential (Tamayo-Molano, 2007). This disease reduces quality and degrades production to second-rate fruit in the domestic market. The incidence of the pathogen can be reduced through a management system based on a combination of effective and timely fungicide sprays, as well as periodic removal of diseased fruit and dead branches within the canopy. Thus, the economic losses caused by scab could be reduced (Hartill, 1991). To avoid over-application of chemical products, growers require alternatives to ensure delivery of quality fruit (Korsten, 2006). Peel color in fruits is a quality attribute that influences consumer preference, and induces the expectation of flavor, taste and palatability (Wadhwa & Capaldi-Phillips, 2014). Therefore, it is useful to construct color scales using instruments such as colorimeters, digital image processing and other objective color determination techniques to reduce the subjectivity of operators in the field or postharvest (Castro-Camacho et al., 2013).

Sphaceloma perseae infection results in a brown, corky-looking, irregularly-shaped lesion in the epidermis, but does not produce pulp rot (Tamayo-Molano, 2007). There are few epidemiological works describing the temporal progress of avocado scab, such as graphical scans (Marroquín-Pimentel, 1999) and some spatial distribution models of thrips (Rivera-Martínez et al., 2017). The incidence of the disease (97 %) has been directly related to the incidence of thrips damage (90 %) (Ávila-Quezada et al., 2003). However, the exact damage caused by this disease is not known (Ávila-Quezada et al., 2002).

In Mexico, avocado scab is commonly controlled with continuous applications of agrochemicals, without prior knowledge of the epidemiology of the pathosystem, which negatively affects the environment and the grower's economic situation (Ávila & Marroquín, 2007). Visual scales are a useful tool to evaluate the severity of plant diseases. These have established classes with a 0-10 rating scale, where 0 is a healthy fruit (any other organ or plant) and 10 is a completely damaged organ (Everett, 1999). Currently, technological tools are available that can help determine the percentage of damage to a crop, such as digital image analysis, information analysis through statistical packages, and different color formats, among others. One color format in which an image can be analyzed is the CIE L*a*b* format, which consists of three important data (Iñiguez et al., 1995): L (lightness), "a" (green to red) and "b" (blue to yellow). Bai et al. (2013) used the CIE L*a*b* format in their research, and by taking illumination into account in the learning stage of the model they

Introducción

La roña del fruto es ocasionada por el hongo *Sphaceloma perseae*, y es uno de los principales problemas fitosanitarios del aguacate (*Persea americana* Mill.), el cual dificulta la exportación (Tamayo-Molano, 2007). Esta enfermedad disminuye la calidad y demerita la producción a fruta de segunda en el mercado nacional. La incidencia del patógeno se puede reducir mediante un sistema de manejo basado en la combinación de aspersiones de fungicidas efectivos y oportunos, así como la eliminación periódica de frutos enfermos y ramas muertas dentro del dosel. Por lo tanto, las pérdidas económicas causadas por la roña podrían disminuir (Hartill, 1991). Para evitar el exceso de la aplicación de productos químicos, los productores requieren alternativas para asegurar la entrega de frutos con calidad (Korsten, 2006). El color de la cáscara en los frutos es un atributo de calidad que influye en la preferencia del consumidor, e induce la expectativa de sabor, gusto y palatabilidad (Wadhwa & Capaldi-Phillips, 2014). Por ello, es útil construir escalas de colores utilizando instrumentos como colorímetros, procesamiento digital de imágenes y otras técnicas de determinación objetivas de color para reducir la subjetividad de los operadores en campo o postcosecha (Castro-Camacho et al., 2013).

La infección por *Sphaceloma perseae* da como resultado una lesión de color café de aspecto corchoso y de forma irregular en la epidermis, pero no produce pudrición de la pulpa (Tamayo-Molano, 2007). Existen pocos trabajos epidemiológicos que describen el progreso temporal de la roña del aguacate, como exploraciones gráficas (Marroquín-Pimentel, 1999) y algunos modelos de distribución espacial de trips (Rivera-Martínez et al., 2017). La incidencia de la enfermedad (97 %) se ha relacionado directamente con la incidencia del daño por trips (90 %) (Ávila-Quezada et al., 2003). Sin embargo, no se conoce con exactitud el daño que presenta esta enfermedad (Ávila-Quezada et al., 2002).

En México, la roña del aguacate se controla comúnmente con aplicaciones continuas de agroquímicos, sin previo conocimiento de la epidemiología del patosistema, lo cual afecta negativamente al ambiente y a la economía del productor (Ávila & Marroquín, 2007). Las escalas visuales son una herramienta útil para evaluar la severidad de las enfermedades en plantas. Estas poseen clases establecidas con una amplitud que puede ir de 0 a 10, donde 0 es fruto sano (cualquier otro órgano o planta) y 10 órgano completamente dañado (Everett, 1999). Actualmente, se dispone de herramientas tecnológicas que pueden ayudar a determinar el porcentaje de daño en un cultivo, como análisis digital de imágenes, análisis de información por medio de paquetes estadísticos, formatos de color diferentes, entre otras. Un formato de color en el que se puede analizar una imagen es

proposed, they discovered that by analyzing colors in images with non-complex textures a reasonable classification of crops with respect to soil is obtained.

In this study, the CIE L*a*b* color format was used to evaluate the percentage of damage due to scab, considering that scab has a brown color in different shades. Therefore, the objective of this study was to provide an innovative and reliable tool, through image analysis, to evaluate the surface damaged by scab in avocado fruit. This will help to make timely decisions and achieve effective management. The images used contained Hass avocado fruits at developmental stages five and six (Ávila-Quezada et al., 2005).

Materials and methods

The strategy to identify color, proposed in this work, consisted of sampling pixels of one color (with different shades of the same color) from an image (Figure 1 and Figure 2). The RGB data of the sampled

el formato CIE L*a*b*, el cual consta de tres datos importantes (Iñiguez et al., 1995): L (luminosidad), “a” (de verde a rojo) y “b” (de azul a amarillo). Bai et al. (2013) utilizaron el formato CIE L*a*b* en su investigación, y al tomar en cuenta la iluminación en la etapa de aprendizaje del modelo que propusieron descubrieron que al analizar colores en imágenes con texturas no complejas se obtiene una razonable clasificación de cultivos con respecto al suelo.

En este estudio se utilizó el formato de color CIE L*a*b* para evaluar el porcentaje de daño por roña, esto considerando que la roña tiene un color café en distintas tonalidades. Por lo tanto, el objetivo del presente estudio fue aportar una herramienta innovadora y confiable, mediante análisis de imágenes, para evaluar la superficie dañada por roña en frutos de aguacate. Esto ayudará a tomar decisiones oportunas y lograr un manejo eficaz. Las imágenes que se utilizaron contenían frutos de aguacate cv. Hass en etapa de desarrollo cinco y seis (Ávila-Quezada et al., 2005).



Figure 1. Section image of a tomato plant.

Figura 1. Imagen de sección de una planta de jitomate.

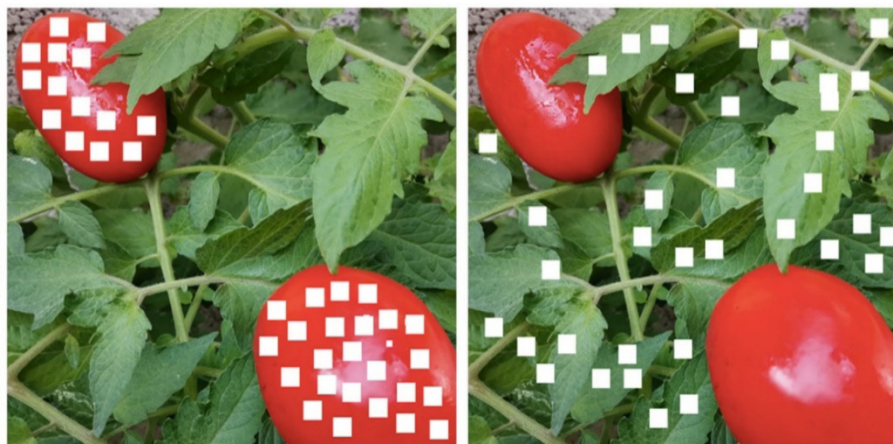


Figure 2. Sample of target pixels: red (left) and green (right) in a tomato plant.

Figura 2. Muestra de pixeles objetivo: rojos (izquierda) y verdes (derecha), en una planta de jitomate.

pixels were transformed to the CIE L*a*b* format. The CIE L*a*b* coordinates were plotted in a three-dimensional coordinate system. When looking at the plotted points from different angles, it was found that in the “a-b” plane view the pixels corresponding to the same color in different shades were in the same neighborhood of points (Figure 3). The pixel points in the “a-b” plane were plotted again (discarding the value of L) and a polygon was drawn around them (Figure 4). Subsequently, the initial image (Figure 1) was analyzed, pixel by pixel, looking for pixels whose “a, b” coordinates were within the polygon which were then counted (Figure 5). The percentage of green pixels in an image was determined by dividing the number of green pixels found by the total number of pixels in the image.

Materiales y métodos

La estrategia para identificar el color, propuesta en este trabajo, consistió en tomar una muestra de píxeles de un color (con diferentes tonalidades del mismo color) de una imagen (Figura 1 y Figura 2). Se transformaron los datos RGB de los píxeles muestreados al formato CIE L*a*b*. Las coordenadas CIE L*a*b* se graficaron en un sistema de coordenadas en tres dimensiones. Al observar los puntos graficados desde diferentes ángulos, se encontró que en la vista del plano “a-b” los píxeles correspondientes a un mismo color en diferentes tonalidades se encontraban en un mismo vecindario de puntos (Figura 3). Se graficaron nuevamente los puntos de los píxeles en el plano

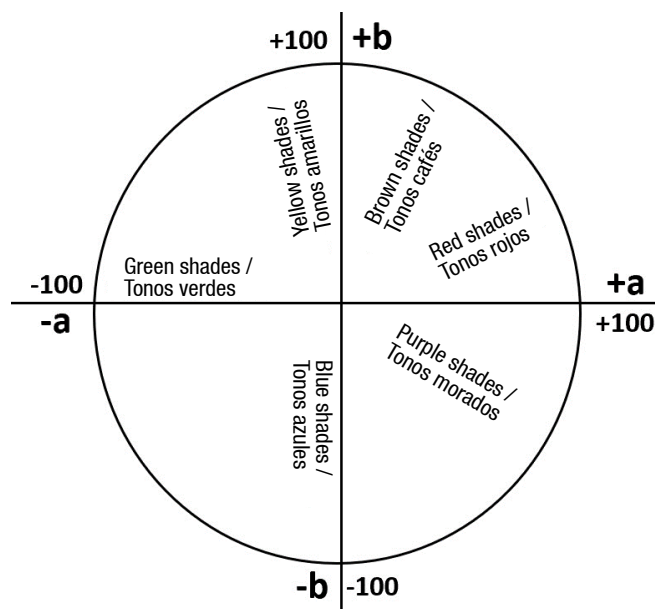


Figure 3. Colors in the CIE L*a*b* format when the L (lightness) value is zero.

Figura 3. Colores en el formato CIE L*a*b* cuando el valor L (luminosidad) es cero.

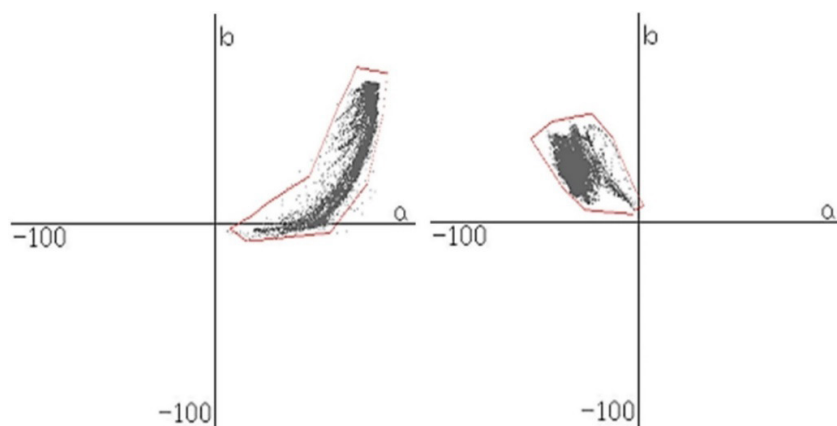


Figure 4. Red (left) and green (right) pixels sampled and plotted on the “a-b” plane (using CIE L*a*b* color format data) with their respective polygons.

Figura 4. Píxeles rojos (izquierda) y verdes (derecha) muestreados y graficados en el plano “a-b” (utilizando los datos del formato de color CIE L*a*b*) con sus respectivos polígonos.



Figure 5. Red pixels changed to white (left), and green pixels changed to white (right).

Figura 5. Píxeles color rojo cambiados a color blanco (izquierda), y píxeles color verde cambiados a color blanco (derecha).

Avocado scab control is carried out when the fruit is green, that is, in the stages prior to ripening. Images of Hass avocado fruit were obtained when the fruit was clearly green. When obtaining photos of avocado fruit in the field, it was not necessary to cut the fruit. Seventy images were obtained with different lighting conditions at different times of the day in orchards in Michoacán, Mexico, because some fruits were exposed to the sun, others were completely shaded, and some were partially shaded. In addition, the photos were taken with the fruit in front of a light blue or white sheet of paper to avoid the leaves of the avocado tree in the image. This ensured that the algorithm recognized only the green of the avocado fruit.

With the images obtained, a sample extraction of only green pixels was made. These pixels were plotted on the “a-b” plane of the CIE L*a*b* color format. Then a polygon was drawn around the projected pixels (Figure 6). The coordinates found for the polygon in “a, b” format are: -25, 80; -14, 67; -7, 26; -4, 14; -5, 0; -10, -6; -20, -8; -50, -10; -71, 5; -85, 32; -79, 65; -58, 75; -31, 81. Pixels that had “a-b” coordinates included within the polygon were green in the subsequent analysis.

To check the correct functioning of the polygon, the number of green pixels (pixels that fell within the polygon) in an image of the avocado fruit was calculated and, as a result, white pixels were placed where green coloring was found (Figure 7). In that image, it was found that there were 42.925 % green pixels. To obtain the percentage of green pixels, the number of green pixels was divided by the total number of pixels in the image, and the result was multiplied by 100. With the above, it was visually found that the polygon had correctly identified all pixels with green coloring.

“a-b” (descartando el valor de L) y se dibujó un polígono alrededor de estos (Figura 4). Posteriormente, se analizó la imagen inicial (Figura 1), pixel por pixel, en busca de píxeles cuyas coordenadas “a, b” estuvieran dentro del polígono y se contabilizaron (Figura 5). El porcentaje de píxeles verdes en una imagen se determinó al dividir el número de píxeles verdes encontrados entre el número total de píxeles de la imagen.

El control para la roña del aguacate se realiza cuando el fruto se encuentra de color verde; es decir, en las etapas previas a la maduración. Se obtuvieron imágenes de frutos de aguacate cv. Hass cuando este se encontraba con un color claramente verde. Al obtener las fotos de frutos de aguacate en campo, no fue necesario cortar el fruto. Se obtuvieron 70 imágenes con diferentes condiciones de iluminación en diferentes horarios del día, en huertos de Michoacán, México, esto debido a que algunos frutos estuvieron expuestos al sol, otros estuvieron completamente en la sombra y algunos estuvieron parcialmente sombreados. Además, se tomaron las fotos con el fruto delante de una hoja de papel de tono azul claro o blanco para evitar hojas del árbol de aguacate en la imagen. Con esto se aseguró que el algoritmo reconociera únicamente el verde del fruto de aguacate.

Con las imágenes obtenidas, se hizo una extracción de muestras de píxeles únicamente verdes. Estos píxeles se graficaron en el plano “a-b” del formato de color CIE L*a*b*. Después, se procedió a dibujar un polígono alrededor de los píxeles proyectados (Figura 6). Las coordenadas encontradas del polígono en formato “a, b” son: -25, 80; -14, 67; -7, 26; -4, 14; -5, 0; -10, -6; -20, -8; -50, -10; -71, 5; -85, 32; -79, 65; -58, 75; -31, 81. Los píxeles que tuvieron coordenadas “a-b” incluidas dentro del polígono fueron verdes en el análisis posterior.

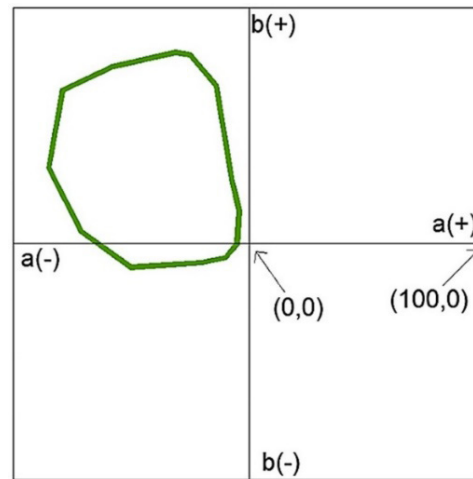


Figure 6. Polygon surrounding all green pixels (sampled from 70 avocado fruit images) plotted in the “a-b” plane of the CIE L*a*b* color format.

Figura 6. Polígono que rodea a todos los píxeles de color verde (muestreados en 70 imágenes de frutos de aguacate) graficado en el plano “a-b” del formato de color CIE L*a*b*.



Figure 7. Original image of an avocado fruit (left) and an image where green pixels were changed to white (right) and counted to estimate the percentage of green pixels in an image. The percentage of green color found in this image was 42.92 %.

Figura 7. Imagen original de un fruto de aguacate (izquierda) e imagen en donde los píxeles color verde se cambiaron a color blanco (derecha) y se contabilizaron para estimar el porcentaje de píxeles verdes en una imagen. El porcentaje de color verde encontrado en esta imagen fue de 42.92 %.

Subsequently, the percentage of scab color was obtained; to do this, the same procedure was performed as with the green color, except that this time only the scab-colored pixels were identified and plotted on the “a-b” plane of the CIE L*a*b* color format (Figure 8). The coordinates of the polygon in “a, b” format were: 7, 50; 16, 49; 25, 47; 30, 30; 26, 22; 20, 16; 12, 10; 5, 8; 3, 13; 0, 18; 4, 34; 4, 49; 5, 51.

To check for correct operation, we visually assessed whether the polygon correctly identified the scab. To do this, in the original image, the pixels found with the scab color were changed to white (Figure 9). In the

Para comprobar el correcto funcionamiento del polígono, se calculó el número de píxeles verdes (píxeles que cayeron dentro del polígono) en una imagen del fruto de aguacate y, como resultado, se colocaron píxeles color blanco en donde se encontró color verde (Figura 7). En dicha imagen se encontró que había 42.925 % de píxeles de color verde. Para obtener el porcentaje de píxeles verdes, se dividió el número de píxeles verdes entre el número total de píxeles en la imagen, y el resultado se multiplicó por 100. Con lo anterior, se encontró visualmente que el polígono había identificado correctamente a todos los píxeles con coloración verde.

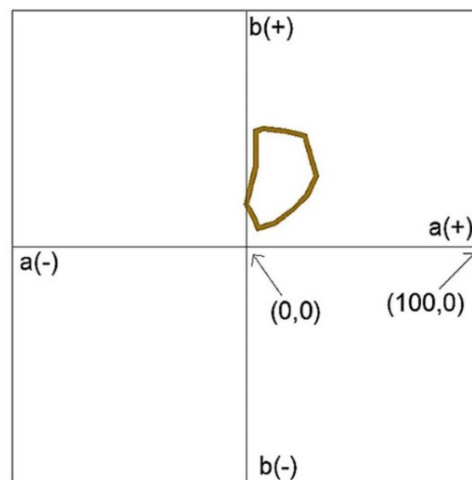


Figure 8. Polygon surrounding all scab-colored pixels (sampled from 70 avocado fruit images) plotted on the “a-b” plane of the CIE L*a*b* color format.

Figura 8. Polígono que rodea a todos los píxeles del color de la roña (muestreados en 70 imágenes de frutos de aguacate) graficado en el plano “a-b” del formato de color CIE L*a*b*.



Figure 9. Original image of an avocado fruit (left) and an image where the pixels found with scab color were changed to white (right) and counted to estimate the percentage of pixels with scab in the image. The percentage of scab color found in this image was 17.37 %.

Figura 9. Imagen original de un fruto de aguacate (izquierda) e imagen en donde los píxeles encontrados con color de la roña se cambiaron a color blanco (derecha) y se contabilizaron para estimar el porcentaje de píxeles con roña en la imagen. El porcentaje del color de la roña encontrado en esta imagen fue de 17.37 %.

image, it was observed that there were 17.37 % of pixels with scab, which was estimated with respect to the total number of pixels in the image.

Because it is necessary to know the percentage of scab with respect to the fruit, it was considered that the sum of the number of green pixels in the entire image, plus the number of scab-colored pixels can be related to the approximate amount of 100 % of pixels that make up the fruit. In addition, the number of scab-colored pixels can be related to the percentage of scab-colored pixels that exist with respect to the fruit. Therefore, by

Posteriormente, se obtuvo el porcentaje de color de la roña; para ello, se realizó el mismo procedimiento que con el color verde, excepto que en esta ocasión se identificaron únicamente los píxeles del color de la roña y se graficaron en el plano “a-b” del formato de color CIE L*a*b* (Figura 8). Las coordenadas del polígono en formato “a, b” fueron: 7, 50; 16, 49; 25, 47; 30, 30; 26, 22; 20, 16; 12, 10; 5, 8; 3, 13; 0, 18; 4, 34; 4, 49; 5, 51.

Para comprobar el correcto funcionamiento, se evaluó de manera visual si el polígono identificó correctamente a la roña. Para ello, en la imagen original, se cambiaron

means of a rule of three, the percentage of scab-colored pixels (percentage of the disease) with respect to the fruit is obtained.

$$P_{scab} = \frac{S \times 100}{G + S}$$

where P_{scab} is the percentage of scab in the fruit, S is the percentage of pixels with scab found with respect to all pixels in the image and G is the percentage of green pixels found with respect to all pixels in the image. Substituting the above values ($S = 17.37\%$ and $G = 42.925\%$) into the equation gives a fruit scab percentage of 28.8 %.

$$P_{scab} = \frac{17.37 \times 100}{42.925 + 17.37} = 28.8\%$$

Subsequently, the percentage of scab on the fruit was calculated with the same images using the diagrammatic scale (Ávila & Marroquín, 2007). All avocado images were digitized in the AutoCAD® program (AUTODESK, 2020), the percentage of scab on the fruit was calculated and the results were compared with the diagrammatic scale. The scab percentages estimated with AutoCAD® (AUTODESK, 2020) were taken as actual or real (observed) data due to the program's high precision.

To analyze the data, the residuals of all scab percentage estimates were first checked for normal behavior; this was done in the R Statistics program (R Development Core Team, 2020) using the *probplot* command of the *e1071* package reported by Meyer et al. (2019). As a next step, the scab percentages obtained with the polygons (abscissae) were plotted against those obtained with AutoCAD® (AUTODESK, 2020) (ordinates). The regression equation of the form $y = \beta_0 + \beta_1(x)$ was found and the coefficient β_0 was analyzed with null hypothesis $\beta_0 = 0$, according to the methodology proposed by Infante-Gil and Zárate-de Lara (1984) to determine if the straight line passes through the coordinate (0, 0). Similarly, the coefficient β_1 was analyzed with null hypothesis $\beta_1 = 1$, also with the methodology of Infante-Gil and Zárate-de Lara (1984) to determine if for each unit increase in the abscissae the ordinates increase by one unit.

Results and discussion

To validate the green and scab color identification polygons of the avocado, the actual percentage of scab on the fruit was obtained. The avocado images were digitized in the AutoCAD® program (AUTODESK, 2020). First, the area of the entire fruit (A_1) was digitized, as shown in the dotted line in Figure 10b, and then the area of the fruit containing scab (A_2) was digitized, as shown in the solid line in Figure 10b. The percentage of scab on the fruit was estimated by dividing the areas A_2/A_1 and multiplying it by 100.

a color blanco los pixeles que encontró con color de la roña (Figura 9). En la imagen se observó que había un 17.37 % de pixeles con roña, lo cual se estimó con respecto al total de pixeles en la imagen.

Debido a que se requiere conocer el porcentaje de roña con respecto al fruto, se consideró que la suma del número de pixeles verdes que hay en toda la imagen, más el número de pixeles color roña, se puede relacionar con la cantidad aproximada al 100 % de pixeles que conforman al fruto. Además, el número de pixeles color roña se puede relacionar con el porcentaje de pixeles color roña que hay con respecto al fruto. Por lo tanto, mediante una regla de tres se obtiene el porcentaje de pixeles color roña (porcentaje de la enfermedad) que hay con respecto al fruto.

$$P_{roña} = \frac{R \times 100}{V + R}$$

donde $P_{roña}$ es el porcentaje de roña en el fruto, R es el porcentaje de pixeles con roña encontrados con respecto a todos los pixeles de la imagen y V es el porcentaje de pixeles verdes encontrados con respecto a todos los pixeles en la imagen. Al sustituir en la ecuación los valores mencionados anteriormente ($R = 17.37\%$ y $V = 42.925\%$) se obtiene un porcentaje de roña en el fruto de 28.8 %.

$$P_{roña} = \frac{17.37 \times 100}{42.925 + 17.37} = 28.8\%$$

Posteriormente, se calculó el porcentaje de roña en el fruto con las mismas imágenes mediante la escala diagramática (Ávila & Marroquín, 2007). Todas las imágenes de aguacate se digitalizaron en el programa AutoCAD® (AUTODESK, 2020), se calculó el porcentaje de roña en el fruto y se compararon los resultados con la escala diagramática. Los porcentajes de roña estimados con AutoCAD® (AUTODESK, 2020) se tomaron como datos reales (observados) debido a la alta precisión del programa.

Para analizar los datos, primero se comprobó que los residuales de todas las estimaciones de porcentajes de roña tuvieran comportamiento normal, esto se realizó en el programa R Statistics (R Development Core Team, 2020), utilizando el comando *probplot* de la paquetería *e1071* reportada por Meyer et al. (2019). Como siguiente paso, se graficaron los porcentajes de roña obtenidos con los polígonos (abscisas), contra los obtenido con AutoCAD® (AUTODESK, 2020) (ordenadas). Se encontró la ecuación de regresión de la forma $y = \beta_0 + \beta_1(x)$ y se analizó el coeficiente β_0 con hipótesis nula $\beta_0 = 0$, de acuerdo con la metodología propuesta por Infante-Gil y Zárate-de Lara (1984) para determinar si la recta pasa por la coordenada (0, 0). De igual manera, se analizó el coeficiente β_1 con hipótesis nula $\beta_1 = 1$, también con la metodología de Infante-Gil y Zárate-de Lara (1984) para

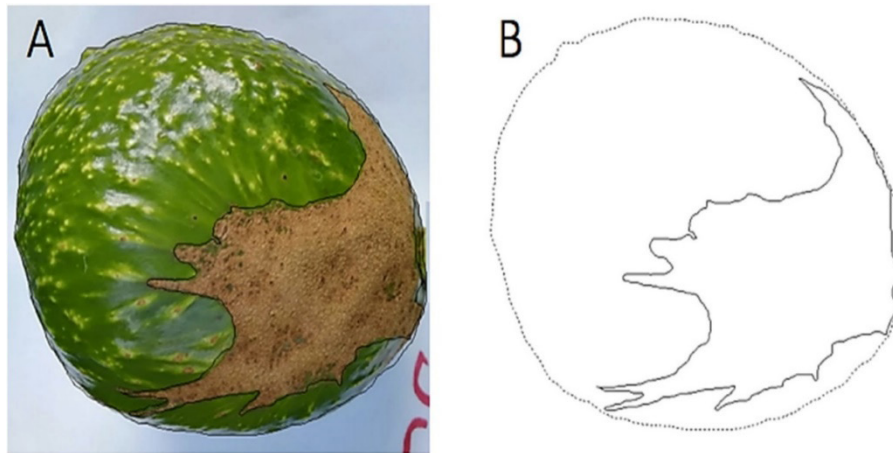


Figure 10. a) Photo of avocado with scab and b) area of the entire fruit (dotted line) and area with scab (solid line) digitalized in AutoCAD®.

Figura 10. a) Foto de aguacate con roña y b) área de todo el fruto (línea punteada) y área con roña (línea continua) digitalizadas en AutoCAD®.

At the visual level, it was observed that the data showed a linear trend when comparing the observed scab percentages with those estimated by the proposed method (Figure 11), and the equation of the straight line with intersection at the origin was obtained ($r^2 = 0.867$).

A first linear regression was performed to obtain the equation of the straight line of the form $y = \beta_0 + \beta_1(x)$. The β_0 value obtained was small compared to the scale of the graph. Because of this, a hypothesis test was carried out on the coefficient β_0 , where the null hypothesis indicated that $\beta_0 = 0$ and the alternate hypothesis indicated that $\beta_0 \neq 0$, with a value of $\alpha = 0.01$. The test statistic ($t_0 = 2.44$) was calculated and compared with the table statistic ($t_{\alpha/2,68} = 2.6479$), which concluded that $\beta_0 = 0$. Therefore, the value of β_0 is not necessary in the model.

On the other hand, a hypothesis test was performed on the coefficient β_1 of the straight line $y = \beta_0 + \beta_1(x)$, where the null hypothesis indicated that $\beta_1 = 1$ and the alternate hypothesis indicated that $\beta_1 \neq 1$. Likewise, the test statistic ($t_0 = -0.17$) was calculated and compared with the table statistic ($t_{\alpha/2,68} = 2.6479$), with which the null hypothesis of $\beta_1 = 1$. With this value it can be said that the β_1 of the equation of the straight line in Figure 11 is equal to 1. In addition, it is statistically proven that the equation of the straight line indicates that for each unit increase in the value of the scab percentage found with the help of the polygon, the actual value of the scab percentage in the fruit obtained with the Autocad® program (AUTODESK, 2020) increases by one unit.

The regression of the equation of the straight line of the actual scab percentage with the scab percentage

determinar si por cada unidad que incrementan las abscisas incrementa la unidad las ordenadas.

Resultados y discusión

Para validar los polígonos de identificación de color verde y roña del aguacate, se obtuvo el porcentaje real de roña del fruto. Se digitalizaron las imágenes de aguacate en el programa AutoCAD® (AUTODESK, 2020). Primero se digitalizó el área de todo el fruto (A_1), como se observa en la línea punteada de la Figura 10b y, posteriormente, se digitalizó el área del fruto que contenía roña (A_2), como se muestra en la línea continua de la Figura 10b. El porcentaje de roña en el fruto se estimó al dividir las áreas A_2/A_1 y multiplicarlo por 100.

A nivel visual, se observó que los datos mostraron una tendencia lineal al comparar los porcentajes de roña observados con los estimados por el método propuesto (Figura 11), y se obtuvo la ecuación de la recta con intersección en el origen ($r^2 = 0.867$).

Se realizó una primera regresión lineal para obtener la ecuación de la recta de la forma $y = \beta_0 + \beta_1(x)$. El valor de β_0 obtenido fue pequeño en comparación con la escala de la gráfica. Debido a esto, al coeficiente β_0 se le hizo una prueba de hipótesis, en donde la hipótesis nula indicaba que $\beta_0 = 0$ y la hipótesis alterna indicaba que $\beta_0 \neq 0$, con un valor de $\alpha = 0.01$. Se calculó el estadístico de prueba ($t_0 = 2.44$) y se comparó con el estadístico de tablas ($t_{\alpha/2,68} = 2.6479$), con lo cual se concluyó que $\beta_0 = 0$. Por lo tanto, el valor de β_0 no es necesario en el modelo.

Por otro lado, al coeficiente β_1 de la recta $y = \beta_0 + \beta_1(x)$ se le hizo una prueba de hipótesis, en donde la hipótesis

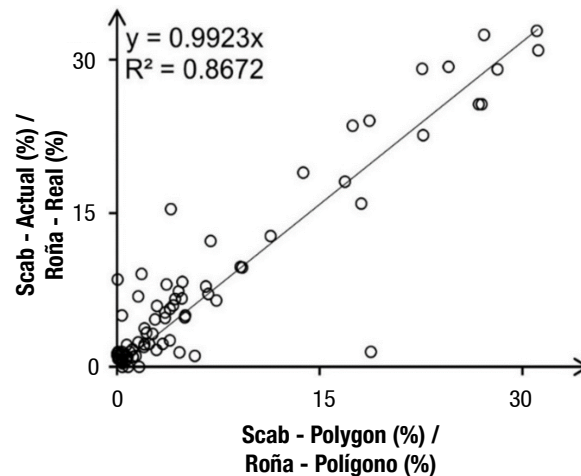


Figure 11. Scab percentage obtained with the polygon-based application of the “a-b” plane of the CIE L*a*b* color format (Scab - Polygon) versus the scab percentage estimated directly with the digitized areas in the image (Scab - Actual).

Figura 11. Porcentaje de roña obtenido con la aplicación basada en polígonos del plano “a-b” del formato de color CIE L*a*b* (Roña - Polígono) contra el porcentaje de roña estimado directamente con las áreas digitalizadas en la imagen (Roña - Real).

found with the diagrammatic scale (Ávila & Marroquín, 2007) was calculated, and a $y = 0.5848x$ was obtained, with an $r^2 = 0.80$.

Ávila and Marroquín (2007) designed a precise, accurate ($R^2 > 0.8$ and $\beta_1 > 0.8$) and reproducible diagrammatic logarithmic scale to evaluate scab in avocado. In the present work, with said diagrammatic scale, an $r^2 = 0.80$ and $\beta_1 = 0.58$ were found. It is likely that, due to the complexity of the symptom patterns and the variable assessment of the evaluator (due to a lack of training), the value of β_1 is sometimes less than 0.8.

These scales are the most commonly used to measure disease severity and are based on the principle of Horsfall and Barrat (1945) and the Weber-Fechner law (Campbell & Madden, 1990). Overestimation of severity levels using logarithmic scales has been reported for *Puccinia horiana* (Barbosa et al., 2006) and for *Ramularia gossypii* (Aquino et al., 2008). Leaves with similar severities, but with a different number of lesions, generate a tendency to overestimate the disease, mainly when the number of lesions is very high and their size is small (Sherwood et al., 1983). The six-class diagrammatic logarithmic scales proposed to assess the severity of *Corynespora cassiicola*-induced leaf and calyx spotting of roselle (Ortega-Acosta et al., 2016) provided good accuracy, precision, and reproducibility in the estimates.

Hock et al. (1992) developed a scale also based on the logarithmic principle for the field evaluation of the severity of the tarspot disease complex of maize; however, the precision values obtained were

nula indicó que $\beta_1 = 1$ y la hipótesis alterna indicaba que $\beta_1 \neq 1$. De igual manera, se calculó el estadístico de prueba ($t_0 = -0.17$) y se comparó con el estadístico de tablas ($t_{\alpha/2, 68} = 2.6479$), con lo cual se aceptó la hipótesis nula de $\beta_1 = 1$. Con este valor se puede decir que el β_1 de la ecuación de la recta de la Figura 11 equivale a 1. Además, se comprueba estadísticamente que la ecuación de la recta indica que por cada unidad incrementada en el valor del porcentaje de roña encontrado con ayuda del polígono se incrementa en una unidad el valor real del porcentaje de roña en el fruto obtenido con el programa Autocad® (AUTODESK, 2020).

Se calculó la regresión de la ecuación de la recta del porcentaje de roña real con el porcentaje de roña encontrado con la escala diagramática (Ávila & Marroquín, 2007), y se obtuvo una $y = 0.5848x$, con una $r^2 = 0.80$.

Ávila y Marroquín (2007) diseñaron una escala logarítmica diagramática precisa, exacta ($R^2 > 0.8$ y $\beta_1 > 0.8$) y reproducible para evaluar roña en aguacate. En el presente trabajo, con dicha escala diagramática, se encontró un $r^2 = 0.80$ y $\beta_1 = 0.58$. Es probable que, debido a la complejidad de los patrones de síntomas, y por la apreciación variable del evaluador (por falta de entrenamiento), el valor de β_1 en ocasiones sea menor a 0.8.

Estas escalas son las más utilizadas para medir la intensidad de la enfermedad, y se basan en el principio de Horsfall y Barrat (1945) y en la ley de Weber-Fechner (Campbell & Madden, 1990). La sobreestimación de los niveles de severidad utilizando escalas logarítmicas

not adequate due to the complicated quantification system and the high number of classes with which it was developed. Michereff et al. (2006) determined suitable precision values in a first evaluation with the use of scales. Tovar-Soto et al. (2002) had to familiarize the evaluators to increase the precision and accuracy values (Hernández-Ramos & Sandoval-Islas, 2015). Due to the use of numerous scales and varying accuracy, some authors such as Bai et al. (2013) and Yakushev and Kanash (2016) used computer programs to identify color regions in crops using data transformed to the CIE L*a*b color format. Bai et al. (2013) obtained 87.2 % in the performance of the model they proposed with the CIE L*a*b color format. In the present work, an r^2 of 0.867 was found, indicating that a performance similar to that of Bai et al. (2013) was obtained.

The coefficient of determination obtained with the polygon is acceptable (> 0.8), which shows that precision can be obtained with the use of this measurement strategy. When comparing the actual scab percentages (obtained with Autocad®) with those obtained with the polygon method, an $r^2 = 0.867$ and a slope of 0.992 were obtained. Therefore, the polygon allowed the assessment of the disease with precision, accuracy and reproducibility. Accuracy is the closeness of an estimated value to the actual value, while precision is the variation or repeatability associated with an estimate, and reproducibility is the absence of variation in estimates when different assessors quantify the same characteristic (Nascimento et al., 2005).

The polygons found cover most of the green and scab-colored variations; therefore, it is recommended to use the same polygon coordinates to determine the scab percentage in any image with a green avocado fruit and a white or light blue background. If the outside of the avocado fruit was not green, an adjustment to the polygon would have to be made; therefore, it is always recommended to visually check the estimate in at least one image.

Conclusions

The method proposed in this study may be useful in research where clearly different colors need to be identified.

With the method proposed in this study, based on polygons, it was found that the equation $y = 0.9923x$, with an $r^2 = 0.8672$, proves that this method performs better than the diagrammatic scale.

The results generated will support future monitoring of the disease in a novel way.

End of English version

ha sido reportada para *Puccinia horiana* (Barbosa et al., 2006) y para *Ramularia gossypii* (Aquino et al., 2008). Las hojas con severidades similares, pero con un número diferente de lesiones, generan una tendencia a sobreestimar la enfermedad, principalmente cuando el número de lesiones es muy alto y su tamaño pequeño (Sherwood et al., 1983). Las escalas logarítmicas diagramáticas de seis clases propuestas para evaluar la severidad del manchado de hojas y cálices de jamaica, inducido por *Corynespora cassiicola* (Ortega-Acosta et al., 2016), proporcionaron buena exactitud, precisión y reproducibilidad en las estimaciones.

Hock et al. (1992) desarrollaron una escala basada también en el principio logarítmico para la evaluación en campo de la severidad del complejo mancha de asfalto del maíz; sin embargo, los valores de precisión obtenidos no resultaron adecuados, esto debido al complicado sistema de cuantificación y al número elevado de clases con el cual fue desarrollada. Michereff et al. (2006) determinaron valores de precisión adecuados en una primera evaluación con el uso de escalas. Tovar-Soto et al. (2002) tuvieron que familiarizar a los evaluadores para incrementar los valores de precisión y exactitud (Hernández-Ramos & Sandoval-Islas, 2015). Debido al uso de numerosas escalas y exactitud diversa, algunos autores como Bai et al. (2013) y Yakushev y Kanash (2016) utilizaron programas de cómputo para identificar regiones de colores en cultivos utilizando datos transformados al formato de color CIE L*a*b. Bai et al. (2013) obtuvieron 87.2 % en el desempeño del modelo que propusieron con el formato de color CIE L*a*b*. En el presente trabajo se encontró una r^2 de 0.867, lo cual indica que se obtuvo un desempeño similar al de Bai et al. (2013).

El coeficiente de determinación obtenido con el polígono es aceptable (> 0.8), lo cual demuestra que se puede obtener precisión con el uso de esta estrategia de medición. Al comparar los porcentajes reales de roña (obtenidos con Autocad®) con los obtenidos con el método de polígonos, se obtuvo una $r^2 = 0.867$ y una pendiente de 0.992. Por lo tanto, el polígono permitió la evaluación de la enfermedad con precisión, exactitud y reproductibilidad. La exactitud es la proximidad que tiene un valor estimado con el valor real, mientras que la precisión es la variación o repetitividad asociadas con una estimación, y la reproducibilidad es la ausencia de variación en las estimaciones cuando diversos evaluadores cuantifican la misma característica (Nascimento et al., 2005).

Los polígonos encontrados abarcan la mayoría de las variaciones de color verde y color de la roña; por lo tanto, se recomienda usar las mismas coordenadas de los polígonos para determinar el porcentaje de roña en cualquier imagen con un fruto de aguacate verde y

References / Referencias

- Aquino, L. A., Berger, P. G., Rodríguez, F. A., Zambolim, L., Hernández, F. R., & Miranda, L. M. (2008). Elaboração e validação de escala diagramática para quantificação da mancha de ramularia do algodoeiro. *Summa Phytopathologica*, 34(4), 361-363. <https://doi.org/10.1590/S0100-54052008000400012>
- AUTODESK, Ink. (2020) *Software documentation for AutoCAD®*. <https://www.autodesk.mx/products?page=2>
- Ávila, Q. G., & Marroquín, P. J. (2007). Enfermedades: la roña del aguacate. In Téliz, D., & Mora, A. (Eds), *El aguacate y su manejo integrado* (pp. 173-208). Mundi – Prensa.
- Ávila-Quezada, G. D., Téliz-Ortiz, D., González-Hernández, H., Vaquera-Huerta, H., Tijerina-Chávez, L., Johansen-Naime, R., & Mojica-Guzmán, A. (2002). Dinámica espacio-temporal de la roña (*Elsinoe perseae*), el daño asociado a trips y antracnosis (*Glomerella cingulata*) del aguacate en Michoacán, México. *Revista Mexicana de Fitopatología*, 20(1), 77-87. https://www.researchgate.net/publication/236569648_Dinamica_espacio-temporal_de_rona_dano_asociado_a_trips_y_antracnosis_en_aguacate
- Ávila-Quezada, G. D., Téliz-Ortiz, D., Mora-Aguilera, G., Vaquera-Huerta, H., & Tijerina-Chávez, L. (2003). Spatial and temporal dynamic of scab (*Sphaceloma perseae* Jenk.) on avocado (*Persea americana* Mill.). *Revista Mexicana de Fitopatología*, 21(2), 152-160. <https://www.redalyc.org/pdf/612/61221209.pdf>
- Ávila-Quezada, G. D., Téliz-Ortiz, D., Vaquera-Huerta, H., González-Hernández, H., & Johansen-Naime, R. (2005). Progreso temporal del daño por trips (Insecta: *Thysanoptera*) en aguacate (*Persea americana* Mill.). *Agrociencia*, 39(4), 441-447. <https://www.redalyc.org/pdf/302/30239408.pdf>
- Bai, X. D., Cao, Z. G., Wang, Y., Yu, Z. H., Zhang, X. F., & Li, C. N. (2013). Crop segmentation from images by morphology modeling in the CIE L* a* b* color space. *Computers and Electronics in Agriculture*, 99, 21-34. <https://doi.org/10.1016/j.compag.2013.08.022>
- Barbosa, M. A., Michereff, S. J., & Mora, A. G. (2006). Elaboration and validation of diagrammatic key to evaluate white rust severity in *Chrysanthemum*. *Summa Phytopathologica*, 32(1), 57-62. <https://doi.org/10.1590/S0100-54052006000100008>
- Campbell, C. L., & Madden, L. V. (1990). *Introduction to plant disease epidemiology*. John Wiley & Sons. New York.
- Castro-Camacho, J. K., Cerquera-Peña, N. E., & Gutiérrez-Guzmán, E. (2013). Determinación del color del exocarpio como indicador de desarrollo fisiológico y madurez en la guayaba pera (*Psidium guajava* cv. *Guayaba pera*), utilizando técnicas de procesamiento digital de imágenes. *Revista EIA*, 10(19), 79-89. <https://revistas.eia.edu.co/index.php/reveia/article/view/498>
- Everett, K. R. (1999). Infection of unripe avocado fruit by stem end rot fungi in New Zealand. *Revista Chapingo Serie Horticultura*, 5, 337-339. http://www.avocadosource.com/WAC4/WAC4_p337.pdf
- Hartill, W. F. T. (1991). Post-harvest diseases of avocado fruits in New Zealand. *New Zealand Journal of Crop and Horticultural* fondo blanco o azul claro. En caso de que el exterior del fruto del aguacate no fuera verde, se tendría que hacer un ajuste del polígono; por ello, siempre se recomienda comprobar visualmente la estimación en al menos una imagen.
- Conclusiones**
- El método propuesto en este estudio puede ser útil en investigaciones en las que se requiere identificar colores claramente diferentes.
- Con el método propuesto en este estudio, basado en polígonos, se encontró que la ecuación $y = 0.9923x$, con una $r^2 = 0.8672$, comprueba que este método se desempeña mejor que la escala diagramática.
- Los resultados generados apoyarán el monitoreo futuro de la enfermedad de una manera novedosa.
- Fin de la versión en español**
- Science, 19(3), 297-304. <https://doi.org/10.1080/01140671.1991.10421814>
- Hernández-Ramos, L., & Sandoval-Islas, J. S. (2015). Escala diagramática de severidad para el complejo mancha de asfalto del maíz. *Revista Mexicana de Fitopatología*, 33(1), 95-103. <https://www.redalyc.org/pdf/612/61240687006.pdf>
- Hock, J., Kranz, J., & Renfro, B. L. (1992). Tests of standard diagrams for field use in assessing the tarspot disease complex of maize (*Zea mays*). *Tropical Pest Management*, 38(3), 314-318. <https://doi.org/10.1080/09670879209371716>
- Horsfall, J. G., & Barratt, W. R. (1945). An improved grading system for measuring plant disease. *Phytopathology*, 35, 655. <http://www.garfield.library.upenn.edu/classics1986/A1986A666500001.pdf>
- Infante-Gil, S., & Zárate-de Lara, G. P. (1984). *Métodos estadísticos. Un enfoque interdisciplinario*. Trillas.
- Iñiguez, M., Ortega, A. P., Rosales, A., Orive, R. A., & Puras, P. (1995). Estudio de color de los vinos tintos de la DOC Rioja. *Zubía*, 7, 167-186. <https://dialnet.unirioja.es/servlet/articulo?codigo=110330>
- Korsten, L. (2006). Advances in control of postharvest diseases in tropical fresh produce. *International Journal Postharvest Technology and Innovation*, 1(1), 48-61. <https://doi.org/10.1504/IJPTI.2006.009181>
- Marroquín-Pimentel, F. J. (1999). Factores que favorecen la incidencia de roña (*Sphaceloma perseae* Jenk.) en el cultivo de aguacate (*Persea americana* Mill.) 'Hass' en tres regiones agroclimáticas de Michoacán, México. *Revista Chapingo Serie Horticultura*, 5, 309-312. http://www.avocadosource.com/WAC4/WAC4_p309.pdf
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C. C., Lin, C. C., & Meyer, M. D. (2019). Package 'e1071'. *The R Journal*. <http://sunsite2.icm.edu.pl/pub/unix/math/cran/web/packages/e1071/e1071.pdf>

- Michereff, S. J., Noronha, M. A., Andrade, D. E., Oliveira, E. P., Xavier, F. M., & Moreira, P. A. (2006). Development and validation of diagrammatic key for *Cercospora* leaf spot of sweet pepper. *Summa Phytopathologica*, 32(3), 260-266. <https://doi.org/10.1590/S0100-54052006000300008>
- Nascimento, R. P., Michereff, J. S., de la Rosa, L. R., & Gómez, M. A. (2005). Elaboração e validação de escala diagramática para cancro bacteriano da videira. *Summa Phytopathologica*, 31(1), 59-64. http://sitav.pga.uem.br/download/escala-diagramatica-cancro-bacteriano-videira_ana-r-p-nascimento_2005.pdf
- Ortega-Acosta, S. Á., Velasco-Cruz, C., Hernández-Morales, J., Ochoa-Martínez, D. L., & Hernández-Ruiz, J. (2016). Escalas logarítmicas diagramáticas para evaluar la severidad del manchado de hojas y cálices de jamaica. *Revista Mexicana de Fitopatología*, 34(3), 270-285. <https://doi.org/10.18781/r.mex.fit.1606-6>
- R Development Core Team (2020). *A language and environment for statistical computing, version 3.6.3*. <https://www.r-project.org/>
- Rivera-Martínez, R., Ramírez-Dávila, J. F., Rubí-Arriaga, M., Domínguez-López, A., Acosta-Guadarrama, A. D., & Figueroa-Figueroa, D. K. (2017). Modelización espacial de trips (Insecta: Thysanoptera) en el cultivo de aguacate (*Persea americana*). *Revista Colombiana de Entomología*, 43(2), 131-140. <https://doi.org/10.25100/socolen.v43i2.5933>
- Sherwood, R. T., Berg, C. C., Hoover, M. R., & Zeiders, K. E. (1983). Illusions in visual assessment of *Stagonospora* leaf spot of orchardgrass. *Phytopathology*, 73(2), 173-177. https://www.apsnet.org/publications/phytopathology/backissues/Documents/1983Articles/Phyto73n02_173.PDF
- Tamayo-Molano, P. J. (2007). Enfermedades del aguacate. *Revista Politécnica*, 3(4), 51-70. <https://revistas.elpoli.edu.co/index.php/pol/article/view/62/49>
- Tovar-Soto, A., Hernández-Martínez, M., Cristóbal-Alejo, J., Romero-Hijo, R., & Mora-Aguilera, G. (2002). Escala logarítmica diagramática de severidad de la mancha negra (*Colletotrichum gloeosporioides* Penz.) en Chirimoyo (*Annona cherimola* M.). *Revista Mexicana de Fitopatología*, 20(1), 103-109. <https://www.redalyc.org/pdf/612/61220117.pdf>
- Wadhera, D., & Capaldi-Phillips, E. D. (2014). A review of visual cues associated with food on food acceptance and consumption. *Eating Behaviors*, 15(1), 132-143. <https://doi.org/10.1016/j.eatbeh.2013.11.003>
- Yakushev, V., & Kanash, E. (2016). Evaluation of wheat nitrogen status by colorimetric characteristics of crop canopy presented in digital images. *Agrárinformatika/Journal of Agricultural Informatics*, 7(1), 65-74. <https://doi.org/10.17700/jai.2016.7.1.268>