Australian Journal of Crop Science

AJCS 15(02):224-232 (2021) doi: 10.21475/ajcs.21.15.02.p2735



ISSN:1835-2707

### Remote hyperspectral sensing for the early detection of *Melanaphis sacchari* Zehntner, 1897 (Hemiptera: Aphididae) infestations in sorghum leaves

# Ruth Artemisa Aguilera Hernández<sup>1</sup>, Manuel Darío Salas Araiza<sup>2</sup>, Adriana Saldaña Robles<sup>3</sup>, Alberto Saldaña Robles<sup>3</sup>, Mónica Trejo Durán<sup>4</sup>, Noé Saldaña Robles<sup>\*3</sup>

<sup>1</sup>Graduate Program in Biosciences, Life Science Division, University of Guanajuato Campus Irapuato-Salamanca, Irapuato, Guanajuato 36500, Mexico

<sup>2</sup>Agronomy Department, Life Science Division, University of Guanajuato Campus Irapuato-Salamanca, Irapuato, Guanajuato 36500, Mexico

<sup>3</sup>Agricultural Engineering Department, Life Science Division, University of Guanajuato Campus Irapuato-Salamanca, Irapuato, Guanajuato 36500, Mexico

<sup>4</sup>Multidisciplinary Studies Department, Engineering Division, University of Guanajuato, Campus Irapuato-Salamanca, Yuriria, Guanajuato 38944, Mexico

#### \*Corresponding author: saldanar@ugto.mx

#### Abstract

This paper aims to study the reflectance signature information of infested and non-infested sorghum leaves (*Sorghum vulgare* L.) by sugarcane aphid (*Melanaphis sacchari*) to discriminate infested sorghum. The study treatments were 0 (0 aphids/leaf), 1 (1-20 aphids/leaf), 2 (21-50 aphids/leaf), 3 (> = 51 aphids/leaf), 4 (> = 51 aphids/leaf + visible damage), 5 (abiotic stress) and 6 (> = 51 aphids/leaf + abiotic stress). An Ocean Optics<sup>TM</sup> HR4000 spectrometer was used. The multifactor ANOVA and Kruskal-Wallis tests at 95% confidence indicated that the reflectance at 402.95, 528.43, 658.36, 788.13, and 965.14 nm wavelengths have significant differences between treatments and with the control. Also Kernel Discriminant analysis was carried out and the combination of the wavelengths centered at 788.17 and 965.14 nm allows 70 % of correct classification of treatments. The results indicate that it is possible to detect M. sacchari infested sorghum by using the spectral information of some specific wavelengths. This study may enable the research of an aerial sensor to make recommendation maps of application pesticides.

Keywords: KDA, PLS-DA, Spectral signature, spectrometer, wavelength.

#### Introduction

Agricultural pests are considered an important factor that negatively impacts crop yields. Aphids are among the pests causing major damage to crop. They feed from suctioned sap and excrete honeydew on the leaves, which in combination with the environmental dust create an environment conducive to fungal growth. This reduces the photosynthetic capacity of the plant by blocking the absorption of certain lengths of the electromagnetic spectrum in the region of chlorophyll production (550-650 nm). Therefore, the visible effects of the pest on the crops occur when the infestation exceeds the economic threshold of the pest (50 aphids/leaf/plant) (Bowling et al., 2016). This is necessary to carry out continuous monitoring in the production fields to avoid the spread of the pest on the entire crop.

As it is known, sorghum is the fifth most important cereal in the world by production volume and acreage. Sorghum is used to feed livestock and produced as food for humans. The sugarcane aphid pest (*Melanaphis sacchari*) present in sorghum is of special interest. The pest has the peculiarity to settle down on the underside of the leaves of sorghum, which makes the detection, monitoring, and the control tasks more difficult. By exceeding the economic threshold, *M. sacchari* causes sorghum to change the color of the leaves to a purple color, followed by chlorosis, necrosis, null growth, delayed flowering, and poor grain filling. Also *M. sacchari* causes losses in the quality and yield of the crop (Singh et al., 2004). The detection, monitoring, and control in early stages of infestation are important activities to avoid large qualitative and quantitative losses of cultivation.

Conventional monitoring of sugarcane aphid is based on physical inspection on the growing and surrounding areas, and pest trapping in the field (Buoro and Imamichi, 2018). Monitoring and control of large areas are integrated activities and represent high costs of economic and human resources. Once the presence of the pest in the crop is detected, pesticides are applied uniformly. The environmental impact of traditional control lies mainly in applied pesticides that do not differentiate between pests and beneficial insects. For this reason, insect populations that help the natural control of pests are reduced. On the economic perspective, the uniform application of pesticides implies a greater expense for the producer thus reducing the profits of the crop (Ding and Taylor, 2016). Hence, farmers must have new technologies that allow them to detect pests even before they reach the economic threshold. A promising area to detect pests on large locations and reduce costs is Remote Sensing. With the help of the Global Positioning System (GPS) and Geographic Information Systems (GIS), variable crop and pest information can be obtained without physical contact. Another advantage of Remote Sensing in agriculture is to detect the presence of the pest even before the first signs appear in the early infestation stage (Zhang and Kovacs, 2012). Well known techniques of remote sensing studies are the multispectral and hyperspectral reflectance of the vegetation. The techniques have sensors with 4 to 20 spectral bands and sensors with more than 20 spectral bands, respectively.

Remote Sensing has shown results in crops such as wheat, for spectrometric detection at certain wavelengths that allow the identification of greenbug aphid (Schizaphis graminum Rondani) and Russian wheat aphid (Diuraphis noxia) (Homoptera : Aphididae) infestations (Riedell and Blackmer, 1999; Yang et al., 2005; Mirik et al., 2006a; Mirik et al., 2006b; Elliot et al., 2009; Luo et al., 2011; Mirik et al., 2012). Between the 720 to 740 nm range and central lengths at 694 and 800 nm the pest, S. graminum, can be detected. The D. noxia can be detected in visible and near-infrared regions between 460 -710 nm and 760-935 nm. Also, the spectral reflectance of infested and non-infested leaves by aphids in crops such as cotton and mustard have been studied (Reisig and Godfrey, 2007; Kumar et al., 2013). In cotton crops, the pest was detected at the reflectance of 850 nm, compared to mustard crops that pest was detected at 550-560 nm, 700-1250 nm, and 1950-2450 nm. The latter wavelength range allowed to differentiate the levels of infestation in mustard. In addition, the use of spectral vegetation indices such as NDVI (Eq. 1) (Rouse et al., 1973), are extensively studied to discriminate between infested and non-infested crops.

 $NDVI = rac{R_{NIR} - R_{BR}}{R_{NIR} + R_{BR}}$  (Eq. 1)

Where, NDVI is the normalized difference of vegetation index;  $R_{NIR}$  is the reflectance in the near-infrared; and  $R_{BR}$  is the reflectance in the visible red.

The NDVI results showed variation and temporal inconsistency due to the influence of soil spectral reflectance (Yang et al., 2005; Mirik et al., 2006a; Mirik et al., 2006b; Mirik et al., 2012; Reisig and Godfrey, 2007).

However, spectral indices are not the only way to detect vegetation stress factors through multi or hyperspectral data. In this sense, Ray et al. (2010), prove the application of mathematical methods, such as stepwise discriminant analysis (SDA). They found the optimum bandwidth differs for different wavelength regions for crop.

Remote Sensing has also been used for the detection of the sugarcane aphid (*Melanaphis sacchari*) pest in sorghum crop (Elliot et al., 2015; Backoulou et al., 2015; Stanton et al., 2017; Backoulou et al., 2018a; Backoulou et al., 2018b). These studies use multispectral cameras that are mounted on drones

to obtain aerial images of sorghum crops. When the images are analyzed either by the NDVI spectral index or the Spatial Patterns the areas of pest infestation are obtained in those images. However, there is no previous work on the analysis of the spectral reflectance of infested and non-infested leaves to detect the wavelengths where the *M. sacchari* pest generates changes in the spectral curve of the sorghum crop.

This study aims to differentiate amongst non-infested and infested sorghum leaves by *Melanaphis sacchari* Zehntner, 1897 (Hemiptera: Aphididae) through the information obtained in the hyperspectral reflectance. The application of mathematical methods such as Kernel Discriminant Analysis (KDA) and the Discrimination Analysis by Partial Least Squares (PLS-DA) could demonstrate their functionality in the classification of spectral information, for differentiating between levels of infestation.

#### **Results and Discussion**

## Spectral signature of sorghum leaves infested and non-infested

The average spectral signatures by infestation levels are shown in Figure 1. As it can be observed, there are important differences between the spectral signature of each infestation level. Due to the complexity of the curve, it is not possible to observe a visual relationship between reflectance curves and infestation levels. However, it can be said that the reflectance of healthy leaves (non-infested) in the visible range of the electromagnetic spectrum is close to 18% at 550 nm. While in the near-infrared the reflectance peak occurs at 790 nm with a value of 51%. When the infestation is in the early stages T1 and T2, the reflectance peak at 550 nm was decreased to 17% and 14%, respectively. While at 790 nm the values are 56% and 48%. When the infestation is highest at T3, 60% of reflectance in the near-infrared (790 nm) is obtained, however, there are no signs of stress in the plant. In stage T4, the reflectance at 550 and 790 nm have values of 24% and 52%, respectively. The stages T5 and T6 show greater reflectance at 550 nm with values of 21% and 38% and lower reflectance at 790 nm with values of 45% and 47 %, respectively.

Therefore, the results show the infestation of *M. sacchari* modifies the spectral curve behavior of the sorghum crop in the early stages.

## Spectral sensitivity by T/TO ratio to discriminate between infested and non-infested sorghum leaves

To find the wavelengths that represent important changes of reflectance towards the presence of the pest, the spectral sensitivity T/T0 ratio was calculated (Figure 2).

The resulting sensitivity curve is an easy way to detect important differences in the reflectance of any infestation level in comparison to non-infested level. When the reflectance (Figure 1) of any infestation level is greater than the reflectance of T0, the sensitivity value (Figure 2) is greater than the baseline (T0/T0). On the contrary, the sensitivity value is less than the baseline when the reflectance value is lower than the reflectance of T0.

Figure 2 shows that major changes in sensitivity, for all infestation levels and abiotic factors, occur from the 260 to

740 nm. The spectral sensitivity curves did not show a uniform behavior in any wavelength range that is useful to derive prediction models of the infestation levels of the pest. The spectral sensitivity peak for all treatments was observed near the wavelength of 415 nm. At approximately 520 nm, a peak is found that corresponds to treatments T5 and T6. Near the 640 nm, there is a peak in spectral sensitivity for treatments T4, T5, and T6. Graphs for treatments T1, T2, and T3 have a flatten behavior, which does not clearly show the sensitivity of the evaluated wavelengths. Overall, the results indicate the possibility of differentiating between two groups of treatments T0, T1, T2 and T3; and treatments T4, T5, and T6. The statistical analysis must test this assumption.

#### Coefficient of determination of first derivative reflectance curve at the measured wavelengths and infestation levels

Figure 3 shows the values of  $R^2$  calculated from the first derivative  $(\partial \rho / \partial \lambda)$ . From this data, the highest values of  $R^2$  were located. The highest coefficients of correlation between the wavelength and the aphid infestation were found at 402.95, 528.43, 658.36, 788.13, and 965.14 nm.

#### ANOVA and Kruskal-Wallis of first derivative of reflectance

Based on the five highest  $R^2$  found for the first derivative of levels of infestation and abiotic factors, statistical analyses were performed in the Statgraphics<sup>TM</sup> Centurion program.

According to the methodology described by Ray et al. (2010), the average of 10 nm from the first derivative of the reflectance, with center at 402.95, 528.43, 658.36, 788.13, 965.14 nm, respectively, were obtained. The normality tests of those data were performed. The results suggest the data at 402.95 and 658.36 nm come from a non-normal distribution. Therefore, for spectral reflectance at such wavelengths, the non-parametric Kruskal-Wallis test was performed. On the other hand, the remaining 3 wavelengths came from normal data and thus, a multifactor ANOVA analysis was used.

Table 1 shows the results of ANOVA and Kruskal-Wallis analysis that indicate statistically significant differences (P > 0.95) amongst reflectance values for at least two of the infestation levels or abiotic factors affecting the sorghum leaves.

Table 2 summarizes a detailed comparison of significant differences at the 95% confidence for all infestation levels and abiotic factors. The results show, at 402.95 nm, the reflectance values allow to discriminate TO from T4 through T6. Also, at 788.13 nm, reflectance allows to discriminate T0 from T2 through T5 except leaves from T1 or T6. Also, by using the reflectance at wavelengths of 788.13 and 965.14 nm, 71% of pair combinations of infestation levels and abiotic factors can be discriminated. Thus, by using the reflectance at the five wavelengths, it is possible to discriminate amongst any pair of combinations. The results indicate that the "Infestation Level" factor significantly influences the dependent variable, "The first derivative". The "Replication" factor did not significantly influence the dependent variable. The results indicate the feasibility of the method to detect early stages of sugarcane aphid infestation.

### Discriminant analysis by partial least square of first derivative of reflectance

Figure 4 shows the results of the PLS-DA for the analyzed cases. The results show the comparison of the PLS-DA models using the Prediction Error Sum of Squares (PRESS). The PRESS value begins with a high variation in components 1 and 2 followed by a decrease in variation. Table 3 shows the five components for the model PLS-DA of infestation levels. This supports Figure 4, because components 1 and 2 have the highest value under PRESS Random Prediction  $R^2$ . This model explains approximately 70% of the variation that is consistent with the results presented in Table 2.

#### Kernel discriminant analysis of first derivative of reflectance

The results of the Kernel discriminant analysis (KDA) are presented in Table 4 and 5. In Table 4 the Wilks' Lambda showed that the first and second discriminant functions have the best performance prediction of the infestation level. In addition, the 3 and 4 discriminant functions have the possibility to discriminate and predict the infestation level. The values of Eigenvalue, Relative percentage, and Canonical correlation are high and corroborate with the Wilks' Lambda results. Table 5 shows the model using the first discriminant function centered on the wavelengths of 402.95, 528.43, 658.36, 788.13, and 965.14 nm. This has 71.43% of cases correctly classified. Table 5 is consistent with the results of the multifactor ANOVA, Kruskal-Wallis, and PLS-DA analysis. Therefore, this confirms the feasibility of using hyperspectral data for the detection of stress caused by M. sacchari, along with the early stages of infestation.

#### Final analysis of results

The spectral information from the wavelengths with a width of 10 nm and centered at 402.95, 528.43, 658.36, 788.13 and 965.14 nm, allows the classification of the stress caused to the sorghum crop due to different levels of M. sacchari infestations and abiotic factors. In addition, the wavelengths allowed to differentiate in low infestation levels, before reaching the economic threshold of the pest at 50 aphids per leaf per plant. Previous studies tested multispectral aerial images to detect the sugarcane aphid (Elliot et al., 2015; Backoulou et al., 2015; Stanton et al., 2017; Backoulou et al., 2018a; Backoulou et al., 2018b). The cameras used in those studies do not have the sensors that correspond precisely at the wavelengths to detect and discriminate the infested zones with sugarcane aphids. Also, other research studies have used the NDVI as a classification index, however the results suggest temporal inconsistency in the data generated (Riedell and Blackmer, 1999; Yang et al., 2005; Mirik et al., 2006a; Mirik et al., 2006b; Mirik et al., 2012; Yang et al., 2009; Reisig and Godfrey, 2007; Kumar et al., 2013). Therefore, the present study results offer an alternative solution from spectral indices of vegetation, showing the feasibility of using mathematical methods such as PLS-DA regression and KDA.

On the other hand, the design of a sensor with at least two of the five centered wavelengths (788.13 and 965.14 nm) as previously mentioned may help the producer obtain

Table 1. Summary of	i anova	multifactorial	and	Kruskal-Wallis.	Factors:	Infestation	Level	and	Replication.	Dependent	variable:	First
derivative reflectance	at the gi	ven wavelengt	h.									

Wavelength (nm)	Data type	Statistical Analysis	<i>p</i> -value
402.95	Non-normal	Kruskal-Wallis	*0.000142789
528.43	Normal	Multifactor ANOVA	*<0.05
658.36	Non-normal	Kruskal-Wallis	*0.0000096571
788.13	Normal	Multifactor ANOVA	*0.0044
965.14	Normal	Multifactor ANOVA	*0.0001

\* *p* < 0.5.



Fig 1. Comparison between mean spectral signatures of sorghum leaves for the treatments of infestation levels and control after the MatLab  $^{\text{m}}$  filter.

Table 2. Summary of statistically significant differences between treatments, at 95% confidence.

Waveleng										Infes	tatior	1 leve									
tn (nm)	0-	0-	0-	0-	0-	0-	1-	1-	1-	1-	1-	2-	2-	2-	2-	3-	3-	3-	4-	4-	5-
	1	2	3	4	5	6	2	3	4	5	6	3	4	5	6	4	5	6	5	6	6
402.95	-	-	-	*	*	*	-	-	*	*	*	-	-	-	*	-	-	*	-	*	*
528.43	-	-	-	-	*	*	-	-	-	*	*	-	-	*	*	-	*	-	*	-	*
658.36	-	-	-	*	*	*	-	-	*	*	*	-	*	-	*	-	-	*	-	*	*
788.13	-	*	*	*	*	-	-	*	*	*	-	-	-	-	-	-	-	-	-	-	*
965.14	*	-	-	-	*	-	*	*	*	-	*	*	-	-	-	-	*	*	*	-	-

\* *p* < 0.5



Fig 2. Spectral sensitivity by wavelength between the spectral reflectance of infested leaves and reflectance of the healthy leaf (control).

Table 3. Mode	el PLS-DA for Infes	tation level.			
Component	% Variation o Y	on R-square	Medium square PRESS	Prediction R-square	PRESS Random Prediction R-square
1	70.1271	70.1271	31.5821	69.6326	69.6326
2	0.998862	71.126	32.0157	69.2157	69.2157
3	0.0715524	71.1975	34.2711	67.047	67.047
4	0.0324379	71.23	35.2017	66.1522	66.1522
5	0.00236786	71.2324	35.2652	66.0911	66.0911

*p* < 0.5 Cases: 49, Cross-validation: delete every 2.



Fig 3. Coefficient of determination for the first derivative of the reflectance curve at the measured wavelength and infestation levels.

Fable 4. KDA summary.	Cases: 49, Int	festation levels	s: 7, Discriminant f	unction: 5.
-----------------------	----------------	------------------	----------------------	-------------

Discriminant	Eigenvalue	Relative	Canonical	Wilks'	<i>p</i> -value
function		percentage	correlation	Lambda	
1	11.4116	78.75	0.95887	0.0107392	*<0.05
2	1.67487	11.56	0.79130	0.133291	*<0.05
3	1.03788	7.16	0.71365	0.356536	*<0.05
4	0.335735	2.32	0.50135	0.726578	*0.0369
5	0.0303803	0.21	0.17171	0.970515	0.5334

\**p* < 0.5.



Fig 4. Comparison of the PLS-DA model to determine the number of components by using the PRESS to find the greater variation between levels of the infestation.

Current	Group	% Foretold								
Treatment	Size	Treatment								
		0	1	2	3	4	5	6		
0	7	6	0	1	0	0	0	0		
		(85.7%)	(0.0%)	(14.3%)	(0.0%)	(0.0%)	(0.0%)	(0.0%)		
1	7	1	4	2	0	0	0	0		
		(14.3%)	(57.1%)	(28.6%)	(0.0%)	(0.0%)	(0.0%)	(0.0%)		
2	7	1	1	4	0	1	0	0		
		(14.3%)	(14.3%)	(57.1%)	(0.0%)	(14.3%)	(0.0%)	(0.0%)		
3	7	0	0	0	6	1	0	0		
		(0.0%)	(0.0%)	(0.0%)	(85.7%)	(14.3%)	(0.0%)	(0.0%)		
4	7	0	0	2	2	2	1	0		
		(0.0%)	(0.0%)	(28.6%)	(28.6%)	(28.6%)	(14.2%)	(0.0%)		
5	7	0	0	0	0	1	6	0		
		(0.0%)	(0.0%)	(0.0%)	(0.0%)	(14.3%)	(85.7%)	(0.0%)		
6	7	0	0	0	0	0	0	7		
		(0.0%)	(0.0%)	(0.0%)	(0.0%)	(0.0%)	(0.0%)	(100.0%)		

**Table 5.** KDA results, percentage of correctly classified cases: 71.43%.



Fig 5. Example of regrown sorghum with infestation of sugarcane aphid (*M. sacchari*), which was used to infest the sorghum plants in this study.



Fig 6. Measurement equipment. A. HR-4000 UV-VIS spectrometer; B. Deuterium and Halogen light sources; C. Optical fiber for reflectance; D. 45 ° fiber carrier; E. Spectralon WS-1 Ocean Optics.

The data filter eliminated data before 260 nm and after 990 nm. A simple moving average of 40 data was applied (It was approximately 10nm)

The standard deviation filter removed atypical spectral signatures of more than 2 standard deviations



The software printed the filtered reflectance file

Fig 7. Flowchart for the program in MatLab<sup>™</sup>.

information about the crop's health status related with the behavior and location of the sugarcane aphid.

#### **Materials and Methods**

#### Experiment location and sampling

The experiment was carried out under macro tunnel conditions in the Life Sciences Division, University of Guanajuato, with coordinates 20° 44' 21.28'' N and 101° 19' 36.41'' W. The study was conducted in the Bajio Guanajuatense region, which ranks second nationwide in the production of sorghum (FIRA, 2019).

The sorghum was sown in the spring-summer of 2018 agricultural cycle and used 20 kg/ha of seed, on three areas of 18 m<sup>2</sup> with 2 m row and separation of 0.75 cm between them. The sown depth was 3 cm. The base fertilization was sown with 400 kg/ha of the Nutrimax<sup>TM</sup> NPK physical mix (20-12-14 + micros). The appropriate agronomic activities were performed, except the application of insecticides. A 40 x 26 thread/inch anti-aphids mesh was placed over the crop, to prevent the spread of the pest to the surrounding crop fields. For the infestation, host plants with the presence of the pest, Johnsongrass (*Sorghum halepense*), and regrowth of sorghum were searched and collected (Figure 5).

The infested plants were cut with scissors and placed in a paper bag. Then, placed in an ice cooler under the shade to reduce the movement of aphids (Jiménez, 2015). Infested plants were placed on the cultivation in the macro tunnel so that the aphids migrate towards the standing crop. Five infested plants were placed on the 2 m groove.

Hyperspectral measurements were performed between the phenological stages 4 and 5 of the crops, at 50 to 60 days after emergence. Based on the appearance and number of aphids per leaf, these were classified into six different infestation levels and abiotic factors. The experiment was completely random.

The sugarcane aphid number was estimated according to the following classification, where the aphid threshold population per leaf per plant was 50 (Bowling et al., 2016):

T0 – Leaves without aphid, 0 aphids per leaf.

T1 – Leaves with low aphid density, 1-20 aphids per leaf.

 $\mathsf{T2}-\mathsf{Leaves}$  with a medium density of aphid: 21-50 aphids per leaf.

T3 – Leaves with a high density of aphid: 51 and more aphids per leaf.

T4 – Leaves with honeydew of aphid: 51 and more aphids per leaf, honeydew, and visible damage.

T5 – Leaves with abiotic stress: leaves stressed by water excess, without fertilization and visible damage.

T6 – Leaves with very high aphid density and abiotic stress, 51 or more aphids per leaf and abiotic stress.

The leaves previously classified at the infestation level were cleaned by removing the aphids and placing them in plastic bags. The removed aphids were returned to the macro tunnel. A total of 49 leaves were collected from different plants, which represents 7 replications by treatment.

#### Spectral data measurement

The measurements were taken between October 22 through October 25, 2018. The spectral reflectance curves of the leaves were obtained with an Ocean Optics<sup>™</sup> HR4000 UV-VIS spectrometer (Ocean Optics, Inc.). The spectral resolution of the spectrometer was 0.26 nm in the range of 190 to 1100 nm. The optical fiber for reflectance was used and placed in a 45° fiber carrier with respect to the horizontal plane. The light source was preheated approximately 1 h before starting the measurements. This is due to the Deuterium and Halogen lamps reaching the same temperature thus avoiding spectral noise. The Spectralon WS-1 Ocean Optics<sup>™</sup> (Ocean Optics, Inc.) was used as the reflectance standard (Figure 6). The spectrometer was configured to perform 10 scans of each leaf before giving an average reading. The measurements were made between 1200 and 1700 local time in central Mexico.

Seven replications of each infestation level, including control (T0), were made. Three measurements in the middle of each leaf were performed. A total of 147 signatures of spectral reflectance of vegetation were obtained.

The data was collected and stored in text format with the Spectra Suite program (Ocean Optics, Inc.). The spectral signatures of the leaves were processed and analyzed using Excel<sup>™</sup> (Microsoft, 2019), MatLab<sup>™</sup> R20017a and Statgraphics<sup>™</sup> Centurion XVI.I (Statgraphics Technologies, Inc., The Plains, Virginia).

#### Data analysis

Since the spectrometer noise/signal ratio was high, a filter was developed in MatLab <sup>™</sup> to reduce unwanted noise (Figure 7). After filtering, the data was analyzed by 4 different statistical methods. The methods are described as followed.

### Spectral sensitivity by T/T0 ratio to discriminate between infested and non-infested sorghum leaves

The spectral sensitivity was carried out in MS Excel <sup>™</sup> by dividing spectral reflectance at each wavelength of the spectral curve of each infestation level from T1 through T6, and reflectance of T0 infestation level (Eq. 2).

$$S(\lambda_i) = rac{R_{Tn}(\lambda_i)}{R_{T0}(\lambda_i)}$$
 (Eq. 2)

Where, S ( $\lambda_i$ ) is the spectral sensitivity value at  $\lambda_i$  wavelength; R<sub>Tn</sub> ( $\lambda_i$ ) is the spectral reflectance measured at  $\lambda_i$  wavelength for the n level of infestation (which can be T1, T2, T3, T4, T5, T6); R<sub>T0</sub> ( $\lambda_i$ ) is the reflectance of the non-infested leaves (T0); and i is the i-th position of the wavelength into the spectral curve. The ratio was calculated by using the reflectance average of the seven replications of each infestation treatment and the reflectance average of the seven replications of T0 (Reidell and Blackmer, 1999).

### Coefficient of determination of the first derivative reflectance curves and levels of infestation

The first derivative of the reflectance curve  $(\partial \rho / \partial \lambda)$  was the difference of reflectance between two consecutive wavelengths, which was computed and assigned to the minor wavelength (Toral, 2002).

The coefficient of determination  $(R^2)$ , in this study, represents the proportion of the variance in the levels of infestation that is predictable from the reflectance measured values at a certain wavelength.

The coefficient of determination  $R_i^2$  was calculated according to Eq. 3 as follows:

$$R_{i}^{2} = \frac{\sum_{i=1}^{n} (\hat{T}_{i} - \bar{T})}{\sum_{i=1}^{n} (T_{i} - \bar{T})}$$
 (Eq. 3)

Where,  $R_i^2$  is the coefficient of determination at the i-th wavelength in the reflectance spectral curves;  $\hat{T}_i$  is the estimated level of infestation according to Eqs. 4, 5 and 6;  $\underline{T}$  is the average of infestation levels;  $T_i$  is the i-th infestation level number; and n is the number of measured reflectance values, which also corresponds to the number of wavelength bands in the spectral curves. The eq. 4 allows estimating  $\hat{T}_i$  and comes from the linear correlation of infestation levels (Eq. 5 and Eq. 6), and reflectance values ( $\rho_i$ ) at a given i-th wavelength of the spectral curves.

$$\begin{split} \hat{T}_{i} &= m_{i}\rho_{i} + b_{i} + \varepsilon_{0} \qquad (Eq. 4) \\ m_{i} &= \frac{p\sum_{j=0}^{p}\rho_{j,i}j - \sum_{j=0}^{p}\rho_{j,i}\sum_{j=0}^{p}j}{p\sum_{j=1}^{p}\rho_{j,i}^{2} - \left(\sum_{j=0}^{p}\rho_{j,i}\right)^{2}} (Eq. 5) \\ b_{i} &= \frac{\sum_{j=0}^{p}j - m_{i}\sum_{j=0}^{p}\rho_{j,i}}{p} \end{split}$$

Where,  $m_i$  is the slope of the linear equation;  $b_i$  is the y-axis intercept constant; and  $\varepsilon_0$  is the random error of the model; p is the number of infestation levels;  $\rho_{j,i}$  is the reflectance value at the i-th wavelength in the spectral curve for the j-th infestation level.

#### ANOVA and Kruskal Wallis

For these analyses, five wavelengths were chosen under the consideration to obtain the greatest  $R^2$  values. The average of 10 nm from the first derivative range were estimated and applied towards the analyses (Ray et al., 2010). The normality of the data at the five wavelengths for all infestation levels were verified. Hence, statistical analyses of multifactor ANOVA or Kruskal-Wallis were applied to the reflectance data. Both statistics were evaluated at p < 0.05.

#### Discriminant Analysis by Partial Least Squares regression

PLS-DA regression is a method used when a property of a physical system is related in some way with an opto-electronic signal. In this study, the infestation levels were predicted as a function of reflectance data PLS-DA regression to find the principal wavelengths. The linear model from PLS regression is represented by Eq. 7:

$$\begin{split} y &= \beta_0 + \beta_1 \rho_1 + \beta_2 \rho_2 + \dots + \beta_n \rho_n + \varepsilon_1 \quad (\text{Eq. 7}) \\ \text{Where, } y \text{ is the predicted level of infestation; } \beta_0 \text{ is the constant initial value of the model; } \beta_1 \text{through } \beta_n \text{ are the linear coefficients that relates the reflectance values of the spectral curve; } \rho_1 \text{through } \rho_n, \text{ are the reflectance at the spectral curve; } and \varepsilon_1 \text{ is the random error of the model.} \end{split}$$

The  $\beta$  is estimated by Eq. 8 as follows:  $\hat{\beta} = (P^T P)^{-1} P^T y$  (Eq. 8)

Where,  $\hat{\beta}$  is the matrix of linear coefficients of the model; *P* is the matrix of reflectance measurement values of all infestation levels;  $P^T$  is the transposed of *P*; and *y* is the matrix of infestation levels. More details can be found in Pérez and Narasimhan (2018). Due to time limitation, PLS-DA was carried out for the first derivative of the reflectance at the 402.95, 528.43, 658.36, 788.13, 965.14 nm.

#### Kernel Discriminant Analysis

The KDA constructs a linear combination of the p input variables observed quantitatively. These linear combinations are functions to help discriminate amongst groups. The KDA was implemented to find model predictions on the infestation levels (group or class) for new sets of reflectance data from sorghum leaves. The KDA used in this study is the Linear discriminant analysis (LDA). More information of LDA can be found at Bandos et al., 2009.

#### Conclusion

This study shows the variation in the spectral reflectance of sorghum leaves caused by M. sacchari infestation. The reflectance at wavelengths with 10 nm bandwidth and centered at 402.95, 528.43, 658.36, 788.13 and 965.14 nm have the sensitivity to discriminate amongst infestation levels of sugarcane aphids in sorghum leaves. The statistical PLS-DA regression method achieved 70% of correct classifications of infestation levels with two spectral bands at 788.13 and 965.14 nm. The KDA method also reached this percentage by using the five mentioned spectral bands. This allows us to conclude that the M. sacchari pest causes spectral changes in the leaves of sorghum crops in the early stages of infestation. This suggests the pest can be detected and discriminated remotely to aid producers in locating the pest on the fields, eventually making decisions on pesticides applications. However, this information must be extrapolated to the field with the design and construction of a sensor installed in a drone capable of measuring the mentioned wavelengths.

#### Acknowledgment

Ruth Artemisa Aguilera Hernández is a student from the Graduate Program in Biosciences, University of Guanajuato and received fellowship 560772 from CONACYT.

#### References

- Backoulou GF, Elliott NC, Giles KL, Alves TM, Brewer MJ, Starek M (2018a) Using multispectral imagery to map spatially variable sugarcane aphid1 infestations in sorghum. Southwest Entomol. 43(1):37-44.
- Backoulou GF, Elliott NC, Giles KL, Brewer MJ, Starek M (2018b) Detecting change in a sorghum field infested by sugarcane aphid1. Southwest Entomol. 43(4): 823-833.
- Backoulou GF, Elliott NC, Giles KL, Mirik M (2015) Processed multispectral imagery differentiates wheat crop stress caused by greenbug from other causes. Comput Electron Agric. 115:34-39.
- Bandos TV, Bruzzone L, Camps-Valls G (2009) Classification of hyperspectral images with regularized linear discriminant analysis. IEEE Trans Geosci Remote Sens. 47(3): 862-873.
- Bowling R, Brewer M, Knutson A, Way M, Porter P, Bynum E, Allen C, Villanueva R (2016) Monitoreo de pulgón amarillo en sorgo. Texas A&M AgriLife. United States of America.
- Buoro AB, Imamichi T (2018) US patent application No. 15/479,494.
- Ding W, Taylor G (2016) Automatic moth detection from trap images for pest management. Comput Electron Agric. 123:17-28.
- Elliott N, Mirik M, Yang Z, Jones D, Phoofolo M, Catana V, Giles K, Michels Jr (2009) Airborne remote sensing to detect greenbug stress to wheat. Southwest Entomol. 34:205-211.
- Elliott NC, Backoulou GF, Brewer MJ, Giles KL (2015) NDVI to detect sugarcane aphid injury to grain sorghum. J Econ Entomol. 108(3):1452-1455
- FIRA (2019) Fideicomisos instituidos en relación con la agricultura. Panorama agroalimentario, sorgo 2019. Retrieved on November 11, 2019 from: <a href="https://www.fira.gob.mx/InvYEvalEcon/EvaluacionIF">https://www.fira.gob.mx/InvYEvalEcon/EvaluacionIF</a>
- Jiménez PI (2015) Estudio de las especies de pulgones y sus enemigos naturales en una finca hortofrutícola ecológica. Universidad Politécnica de Valencia, España.
- Kumar J, Vashisth A, Sehgal VK, Gupta VK (2013) Assessment of aphid infestation in mustard by hyperspectral remote sensing. J Indian Soc Remote Sens. 41:83–90. (doi: 10.1007/s12524-012-0207-6).
- Luo J, Wang D, Dong Y, Huang W, Wang J (2011) Developing an aphid damage hyperspectral index for detecting aphid (Hemiptera: Aphididae) damage levels in winter wheat. In: IEEE IGARSS 2011. p. 1744-1747.

- Mirik M, Ansley RJ, Michels GJ, Elliott NC (2012) Spectral vegetation indices selected for quantifying Russian wheat aphid (Diuraphis noxia) feeding damage in wheat (*Triticum aestivum* L.). Precis Agric. 13:501-516.
- Mirik M, Michels GJ, Kassymzhanova-Mirik S, Elliott NC, Bowling R (2006a) Hyperspectral spectrometry as a means to differentiate uninfested and infested winter wheat by greenbug (Hemiptera: Aphididae). J Econ Entomol. 99:1682-1690.
- Mirik M, Michels GJ, Mirik SK, Elliott NC, Catana V (2006b) Spectral sensing of aphid (Hemiptera: Aphididae) density using field spectrometry and radiometry. Turk J Agric For. 30:421-428. (ISSN: 1300011X).
- Pérez DR, Narasimhan G (2018) So you think you can PLS-DA?. bioRxiv. 207225.
- Ray SS, Jain N, Miglani A, Singh JP, Singh AK, Panigrahy S, Parihar JS (2010) Defining optimum spectral narrow bands and bandwidths for agricultural applications. Curr Sci. 98(10):1365-1369. (Retrieved from www.jstor.org/stable/24107491).
- Reisig D, Godfrey L (2007) Spectral response of cotton aphid– (Homoptera: Aphididae) and spider mite– (Acari: Tetranychidae) infested cotton: controlled studies. Environ entomol. 36:1466-1474.
- Riedell WE, Blackmer TM (1999) Leaf reflectance spectra of cereal aphid-damaged wheat. Crop sci. 39:1835-1840.
- Rouse JW, Haas RH, Schell JA, Deering DW (1973) Monitoring vegetation systems in the great plains with ERTS. NASA special publication. 1:309-317.
- Singh BU, Padmaja PG, Seetharama N (2004) Biology and management of the sugarcane aphid, *Melanaphis sacchari* (Zehntner) (Homoptera: Aphididae), in sorghum: a review. Crop Prot. 23(9):739-755.
- Stanton C, Starek M, Elliott N, Brewer M, Maeda M, Chu T (2017) Unmanned aircraft system-derived crop height and normalized difference vegetation index metrics for sorghum yield and aphid stress assessment. J Appl Remote Sens. 11(2):026035.
- Toral M, Lara N, Tapia AE, Torres C, Richter P (2002) Estudio espectral y determinación simultánea de sulfametoxazol y trimetoprim por espectrofotometría derivada digital. Bol Soc Chil Quím. 47(3):241-251.
- Yang Z, Rao MN, Elliott NC, Kindler SD, Popham TW (2005) Using ground-based multispectral radiometry to detect stress in wheat caused by greenbug (Homoptera: Aphididae) infestation. Comput Electron Agric. 47:121-135.
- Yang Z, Rao MN, Elliott NC, Kindler SD, Popham TW (2009) Differentiating stress induced by greenbugs and Russian wheat aphids in wheat using remote sensing. Comput Electron Agric. 67(1):64-70.
- Zhang C, Kovacs JM (2012) The application of small unmanned aerial systems for precision agriculture: a review. Precis Agric. 13(6):693-712.