Braga AF *et al.* (2023)
Notulae Botanicae Horti Agrobotanici Cluj-Napoca
Volume 51, Issue 3, Article number 13231
DOI:10.15835/nbha51313231
Research Article

NIRS and multivariate methods for discrimination of morning glory species at different growth stages

Andreísa F. BRAGA[1], Lívia C. de CARVALHO[2], Diogo P. CORREA da SILVA[3*], Thamires G. ANTUNES[3], Luis CARLOS CUNHA JÚNIOR[3], Gustavo H. de ALMEIDA TEIXEIRA[4], Pedro L. da COSTA AGUIAR ALVES[1]

[1]*Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências Agrárias e Veterinárias, Departamento de Biologia, Campus de Jaboticabal, Via de Acesso Prof. Paulo Donato Castellane s/n - Jaboticabal/SP, 14884-900, Brazil; flores.braga@unesp.br; pl.alves@unesp.br*

[2]*Universidade Estadual de Maringá, Campus Umuarama, Departamento de Tecnologia, Av. Ângelo Moreira da Fonseca, 1800 - Parque Danielle, Umuarama/PR, 87506-370, Brazil; liviaacirino@gmail.com*

[3]*Universidade Federal de Goiás, Escola de Agronomia, Departamento de Engenharia de Alimentos, Goiânia/GO. Universidade Federal de Goiás Rodovia Goiânia-Nova Veneza, Km 0 S/n Campus, Samambaia, Goiânia/GO, 74690-900, Brazil; diogopedrosa@ufg.br (*corresponding author); cunhajunior.l.c@ufg.br; thamiresg.antunes@hotmail.com*

[4]*University of Idaho (U of I), Kimberly Research and Extension Center. 3806 N 3600 E. Kimberly/ID, Zip code: 83341-5082, USA; gteixeira@uidaho.edu*

Abstract

Morning glory species are weeds very common in tropical crops, where they cause direct and indirect damage. The management of these species primarily relies on the application of herbicides, disregarding the growth stage and spatial distribution. Studies addressing new techniques for identifying these species may contribute to the development of proximal sensors for carrying out specific and rational management. Thus, the objective of this work was to use near infrared spectroscopy (NIRS) and multivariate analysis to discriminate two species of morning glory in three growth stages. NIRS spectra were collected from *Ipomoea hederifolia* and *Merremia aegyptia* were collected at three different stages in the spectral range of 4.000 to 10.000 cm^{-1}. PCA and PC-LDA were used to analyze the entire spectrum and specific bands. NIRS associated with PCA and PC-LDA were sufficient to discriminate *I. hederifolia* and *M. aegyptia* species and their growth stages. PCA allowed a proper segregation of stages and species when applied individually PC-LDA correctly classified between 90.93 to 100% of species and stages. The best discrimination results were observed in the NIR spectra ranges from 4.500 to 6.000 cm^{-1} and 4.500 to 6.000 + 6.500 to 7.750 cm^{-1}. This study represents an advance in the research and implementation of NIRS technology to discriminate weed species for the future development of equipment to assist in the adoption and/or performance of a specific management of weeds, capable of contributing to the reduction in the use of herbicides in crops.

Keywords: *Ipomoea hederifolia*; *Merremia aegyptia*; specific management; weeds; precision agriculture

Received: 31 May 2023. Received in revised form: 18 Jul 2023. Accepted: 23 Aug 2023. Published online: 31 Aug 2023.

From Volume 49, Issue 1, 2021, Notulae Botanicae Horti Agrobotanici Cluj-Napoca journal uses article numbers in place of the traditional method of continuous pagination through the volume. The journal will continue to appear quarterly, as before, with four annual numbers.

Introduction

Agriculture has gone through a great technological revolution. Sensors, robots, algorithms and artificial intelligence are words increasingly present in the field (Klerkx *et al.*, 2019). In this context, the main objective of including new technologies in agricultural systems is related to the adoption of management techniques that enable the increase in productivity in a sustainable way (Van Evert *et al.*, 2017).

Weed management is essential to achieve high productivity and is commonly carried out through the use of herbicides, applied over the entire area (Piza *et al.*, 2016; Alves *et al.*, 2018). The reduction in the use of supplies is one of the main goals in technological advances, as they directly impact operating costs and reduce possible damage to the environment (Klerkx *et al.*, 2019). As a result of rapid technological advancements, numerous studies have proposed various alternatives for conducting pesticide applications exclusively on the intended plant, leading to a significant reduction in the quantity of applied inputs. These approaches aim to minimize environmental impact without compromising efficacy (Wang *et al.*, 2007; Hansen *et al.*, 2013; Singh *et al.*, 2020).

Some of these alternatives for optimizing the use of herbicides is the precise identification of the target, that is, of weeds, through sensors, images and spectral data (Wang *et al.*, 2007; Hansen *et al.*, 2013; Souza *et al.*, 2020). The advantage of developing sensors compared to the use of other technologies is related to the information processing time, directly interfering in decision making and control effectiveness (López-Granados *et al.*, 2010; Souza *et al.*, 2020).

There are already commercially available weed sensors, but they only efficiently detect the presence or absence of these species in contrast to soil or straw (Wang *et al.*, 2007; Lopez-Granados *et al.*, 2010; Singh *et al.*, 2020). The development of sensors to discriminate between weed species is possible, but it requires the selection of better bands and calibration studies, which involve the understanding of spectral properties in relation to other physiological factors, such as chlorophyll and water content, which is influenced by the presence or absence of biotic and abiotic stress (Souza *et al.*, 2020).

The Convolvulaceae family comprises very expressive weed species capable of causing direct and indirect damage to crops (Piza *et al.*, 2016; Alves *et al.*, 2018; Ribeiro *et al.*, 2018). Direct losses are related to competition for nutrients, light and water, which can reduce productivity by up to 36% (Bhullar *et al.*, 2012; Pagnoncelli *et al.*, 2017). The indirect losses are related to harvestability; as they are species with climbing habits, their branches intertwine in the plants, compromising the harvester's operating income and, consequently, impacting operating costs (Piza *et al.*, 2016; Alves *et al.*, 2018).

Near infrared spectroscopy (NIRS) associated with chemometric techniques has been applied in agriculture for a long time (Pasquini, 2003; Nieto-Ortega *et al.*, 2023). The main applications involve quality analysis and characterization of foods and soils (Cunha Júnior *et al.*, 2015; Li *at al.*, 2020; Sohn *et al.*, 2021). Recently, studies have explored its potential for discrimination in plant species, for the purposes of cultivar identification, geographic tracking and weed discrimination (Souza *et al.*, 2020; Vitalis *et al.*, 2020; Nugraha *et al.*, 2021).

However, it is understood that the identification of growth stages is as important as discriminating the species, as the stage interferes with the susceptibility of species to herbicides (Ribeiro *et al.*, 2018).

In view of the above, the objective of this work was to use near infrared spectroscopy (NIRS) and multivariate analysis to discriminate two species of morning glory in three growth stages.

Materials and Methods

Samples

The weed species selected for this study were *Ipomoea hederifolia* and *Merremia aegyptia,* both herbaceous, with climbing habits and belonging to the Convolvulaceae family.

The seeding of these species was carried out in three periods of the year, in plastic pots with a capacity of 0.5 L, filled with washed and sieved river sand. After the emergence of the species, thinning was carried out, leaving only one plant per pot. All plants were cultivated in the open air, under coordinates 21°14'38.6" S and 48°17'56.3" O. Fifty pots were cultivated for each species in each of the development stages, totaling 150 pots per species. The pots were irrigated daily with Hoagland and Arnon's (1950) complete nutrient solution. In the first week, the nutrient solution was adjusted to a concentration of 25%, in the second week to 50% and so on consecutively until it reached 100% ionic strength. The planting period was from March to April 2019, and the climate in the region, according to Köppen's classification, is Aw.

On the day of spectra collection, the plants were classified according to their development stage, according to the BBCH-code (Bleiholder *et al*., 1991). All species were in macro-stage 1 (leaf development), however, due to the particularities of development of each species, the micro-stages were as follows: 10, 11 and 12 for *I. hederifolia* and 11, 12 and 14 *for M aegyptia* (Table 1 and Figure 1).

Table 1. Decimal code for classifying plant development stages

BBCH-code	Definition
1	Leaf development
10	Cotyledonary leaves completely unfolded
11	1st leaf / 1st pair of leaves (or composite leaf) unfolded
12	2nd leaf/2nd pair of leaves (or composite leaf) unfolded
13	3rd leaf/3rd pair of leaves (or composite leaf) unfolded
14	4th leaf/4th pair of leaves (or composite leaf) unfolded

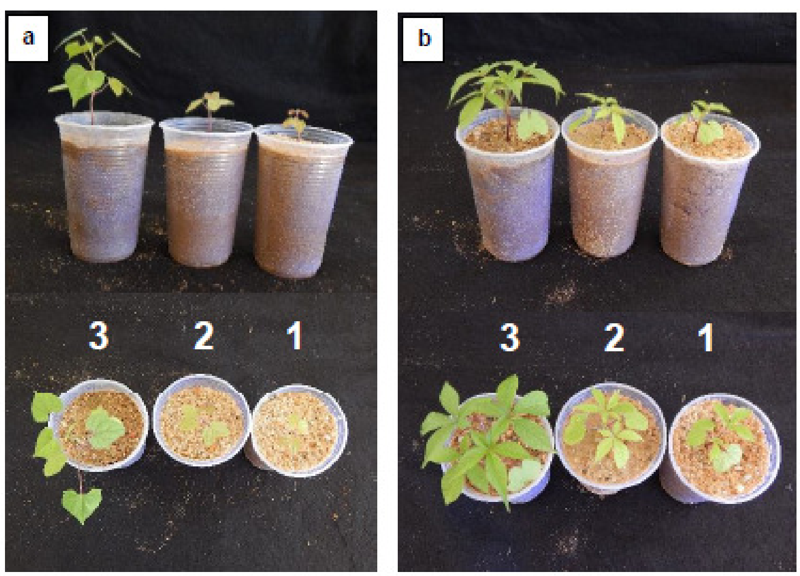

Figure 1. *Ipomoea hederifolia* (a) and *Merremia aegyptia* (b) in 1, 2 and 3 growths stages (corresponding to micro-stages: 10, 11 and 12 in *I. hederifolia* and *11, 12 and 14 for M aegyptia*)

To characterize the stages and species, the water content in the leaves of each species was calculated in each of the micro-stages. For this step, 15 plants of each stage and species were sampled and the fresh mass of leaves was determined. Then, these leaves were placed to dry in an oven with forced air circulation at 40°C for 96 hours, after which the dry mass was determined. The water content in the leaves was expressed as a percentage, being calculated using the equation 1:

Equation 1: WC (%) = ((x-y)/x)*100

Where x represents the fresh mass (g) of the leaves and y represents the dry mass (g) of the leaves. After calculating the water content, the data were submitted to ANOVA {factorial of 3 (stages) x 2 (species) in a completely randomized design} and the means were compared by Tukey test at 5% probability.

Spectra acquisition

NIR spectra were obtained through a NIR Fiber Optic Probe accessory coupled to a FT-NIR Spectrum 100N spectrophotometer® (PerkinElmer®, Shelton, United States of America). The diffuse reflectance spectra were acquired in the form of log (1/R), where R corresponded to the reflectance, in the spectral range of 10.000 to 4.000 cm^{-1} (1.000 to 2.500 nm), with 32 scans, at a spectral resolution of 4 cm^{-1} and with an interval of 2 cm^{-1}. At the time of NIR spectra collection, the ambient temperature was maintained at 23 °C ± 1, remaining constant throughout the spectral acquisition process.

The reading of the spectra was carried out on the adaxial part of the fresh leaves, placed on a black background. The use of black background was adopted to reduce interference from external light input. For each stage, 100 readings of the spectrum (n=300, for each species) were performed.

Chemometric analysis

Spectral data were analysed using the Unscrambler® software version 10.3 (CAMO. Oslo. Norway). The pre-processing used in the spectra was Standard Normal Variate (SNV) plus the 1st Savitzky–Golay derivative (SNV+1SG) grouped into 6 and 6 points, as defined in previous studies (Braga *et al*., 2023).

Principal component analysis (PCA) was performed with the entire NIR spectrum (4.000 to 10.000 cm^{-1}) and in the ranges from 4.500 to 6.000 cm^{-1} and 6.500 to 7.750 cm^{-1}. The spectral ranges were selected according to the PCA loadings values and the peaks of greater reflectance observed in the spectra. To discriminate the stages of species classification models were developed using linear discriminant analysis of principal components (PC-LDA) with raw spectrum and pretreated with SNV+1SG, for all stages and species together.

The separation method within the PC-LDA was the Mahalanobis distance with the use of 10 principal components (PCs) and the random cross validation method, with 10 samples per segment. The spectral database was divided into training (70%) and validation (30%) using the Kennard-Stone algorithm. The performance of the classifiers was evaluated by the percentage of true positives.

Results

Characterization of plant material and near infrared spectra

The largest ranges were selected, with the 1st range spanning from 4500 to 6000 cm^{-1} and the 2nd range spanning from 6500 to 7500 cm^{-1}, as shown in Figure 2.

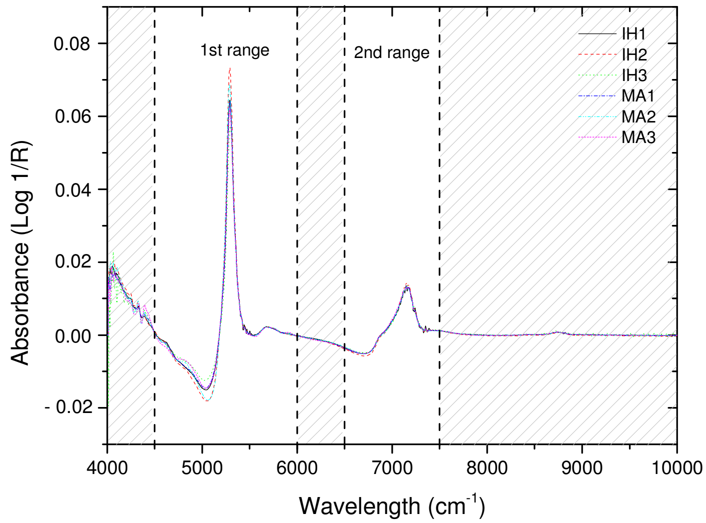

Figure 2. Average absorbance of spectra from 4.000 to 10.000 cm^{-1} with the pre-process

The mean values of water content in *I. hederifolia* in stages 1, 2 and 3 (corresponding to micro-stages: 10, 11 and 12) and *M. aegyptia* in stages 1, 2 and 3 (corresponding to micro-stages: 11, 12 and 14) were significantly different (Table 2). There was a significant difference ($p<0.05$) only for the isolated factors, and the interaction between species and stages factors was not significant ($p = 0.8073$). *I. hederifolia* had higher water content (82.89%) than *M. aegyptia* (80.78%). For the comparison between the stages, the difference observed in the water content was between stage 1 (78.93%), compared to stages 2 (83.33%) and 3 (83.24%).

Table 2. Mean values of water content for *I. hederifolia* and *M. aegyptia* in the three growth stages

Species	Stage 1	Stage 2	Stage 3	Mean Value*
I. hederifolia	80.33	84.25	84.09	82.89^{a}
M. aegyptia	77.53	82.41	82.40	80.78^{b}
Mean values*	78.93^{B}	83.33^{A}	83.24^{A}	81.83

* Means followed by the same letter, uppercase in the row and lowercase in the column, do not differ from each other, at the level of 5% probability.

The average NIR spectra of the species in the three stages in the range of 4,000 to 10,000 cm^{-1} with the SNV+1SG preprocessing showed two main absorbance peaks (Figure 2a). The regions of these peaks coincide with the absorbance peaks of water, which are near 6,800 and 5,500 cm^{-1}, indicating high moisture content in the *Ipomea* samples (Cunha Júnior *et al.*, 2015; Torres *et al.*, 2019).

Principal Component Analysis (PCA)

The analysis of the principal components (PCA) of the whole spectrum (4,000 to 10,000 cm^{-1}) with the pre-processing of SNV+1SG explained 47% of the variables in PC1 and 16% in PC2, that is, 63% in the first two components (Figure 3a). There was a reasonable segregation of groups in the PCA, but in the center of the hyperplane the samples were still overlapped. The samples of *I. hederifolia* segregated the second and first stages, and the samples from the second stage were associated with negative values of PC1 and the first, with positive values. The samples from the third stage were in the center, with positive and negative values for PC1 and PC2.

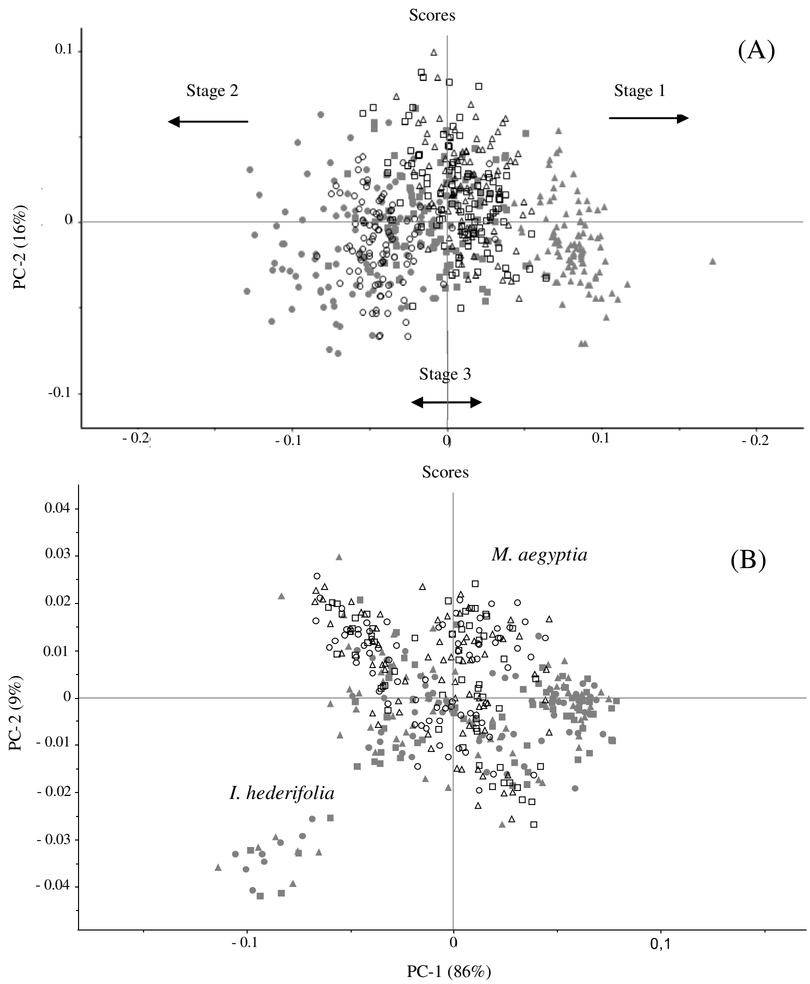

Figure 3. PC 1 and PC 2 values resulting from principal component analysis (PCA) with the entire spectrum (a) and with the two bands (b) for the species *I. hederifolia* (grey) and *M. aegyptia* (black) in three stages of development (Δstage 1; ○: stage 2; □: stage 3).

The species *M. aegyptia* segregated less, but showed the same behavior as *I. hederifolia*, keeping most samples from stage 2 in the negative quadrant of PC1 and from stage 1 in the positive quadrant. The arrows on the graph indicate the orientation of the samples at each stage (Figure 3a). Therefore, the PCA applied to the whole spectrum (4.500 to 10.000 cm^{-1}) was not enough to discriminate the species and stages.

When considering only the two ranges (4.500 to 6.000 cm^{-1} + 6.500 to 7.750 cm^{-1}), the sum of the first two components was more representative for the study of variability (PC1 = 86% and PC2 = 9 %), but the samples also did not segregate (Figure 4b). There was a lot of overlap between species and it was not possible to identify separate groups of stages.

The PCA in the 1st range, from 4.500 to 6.000 cm^{-1} (Figure 4a), discriminated the stages with little overlap and also the species within each stage, except for stage 3, in which the samples of the two species were quite overlapped. PC1 and PC2 explained 87% and 9% of the total spectral variability, respectively. These results were better than those observed in the full spectrum PCA (Figure 2). By analyzing only this region of the spectrum (4.500 to 6.000 cm^{-1}), it was possible to notice differences not only between stages, but also between species. The *M. aegyptia* samples from stages 1 and 2 were mostly in the positive quadrant of PC2 and *I. hederifolia* showed the majority of the negative quadrant of PC2.

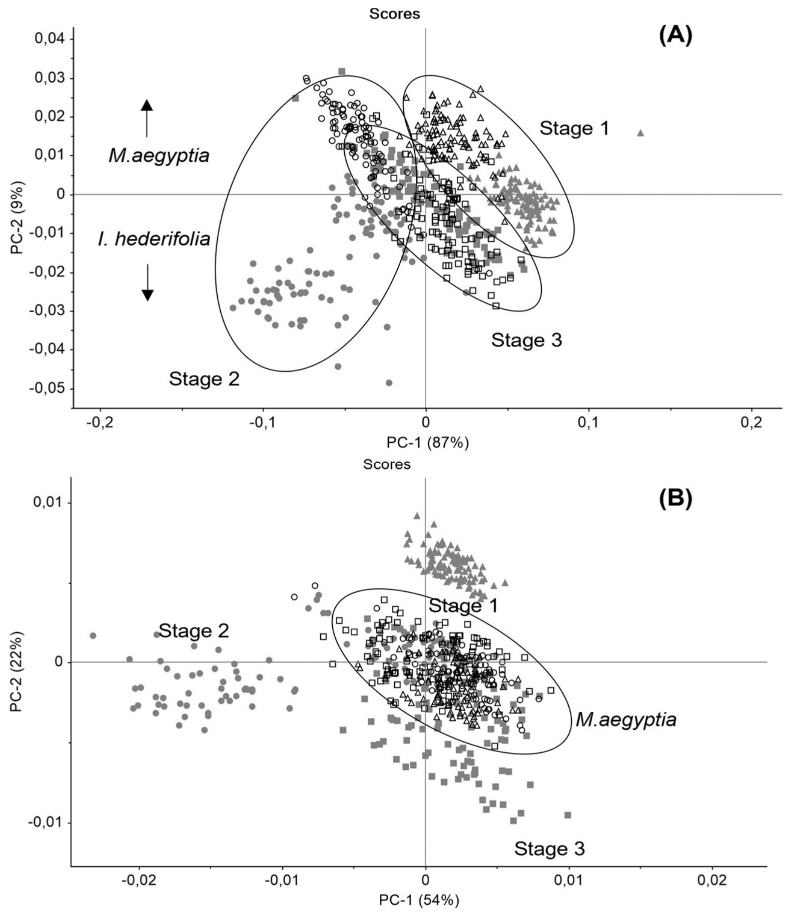

Figure 4. PC 1 and PC 2 values resulting from principal component analysis (PCA) with the 1st band (a) (4.500 a 6.000 cm^{-1}) and with the 2nd band (b) (6.500 a 7.750 cm^{-1}) for the species of *I. hederifolia* (grey) and *M. aegyptia* (black) in three stages of development (Δstage 1; ○: stage 2; □: stage 3)

In the 2nd range, from 6.500 to 7.750 cm^{-1} (Figure 3b), PC1 (54%) and PC2 (22%) were less representative than the 1st range and only segregated some samples of *I. hederifolia* and its stages. *Merremia aegyptia,* in this range, remained at the center of the hyperplane, with a high overlapping of the stages (Figure 3b). Therefore, it is understood that for the study of different stages and species, the 1st band (4.500 to 6.000 cm^{-1} – Figura 4a) is the most suitable.

The PCA for each species in the 1st range, from 4.500 to 6000 cm^{-1} were effective to discriminated between the stages of *I. hederifolia* (Figure 4a) and *M. aegyptia* (Figure 5b) and maintained a high influence of PC1 and PC2. The variability was explained by 93% and 5% for *I. hederifolia,* and by 87% and 11% for *M. aegyptia*, in PC1 and PC2, respectively. The stages in *I. hederifolia* were separated as follows: stage 1, in the positive quadrant of PC1; stage 2, in the negative quadrant of PC1; and stage 3, in the positive quadrant of PC2 (Figure 5a). As for the stages of *M. aegyptia*, the separation of the samples was even more evident: stage 1, positive quadrant for PC1 and negative for PC2; stage 2, negative quadrant of PC1; and stage 3, positive quadrants of PC1 and PC2 (Figure 5b).

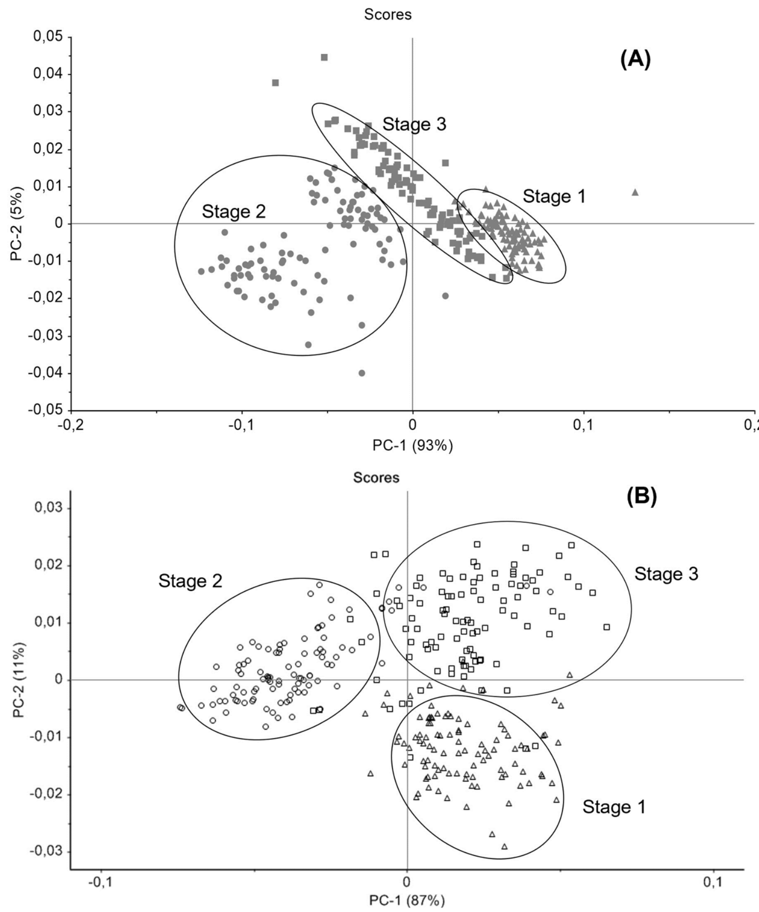

Figure 5. PC 1 and PC 2 values resulting from principal component analysis (PCA) in the 1st band of the spectrum (4.500 to 6.000 cm^{-1}) for the species *I. hederifolia* (a) and *M. aegypti*a (b) in the three stages of growth (Δstage 1; ○: stage 2; □: stage 3)

Linear discriminant analysis (PC-LDA)

The Linear Discriminant Analysis (PC-LDA) applied to the first 10 components of the whole spectrum and of the bands without pre-processing was enough to classify the samples in the six classes (Table 3). The minimum hit percentage was 90.93% and the maximum 100%. The bands with the smallest errors were the peaks added (1st band + 2nd band: 4.500 to 6.000 + 6.500 to 7.750 cm^{-1}) and the 1st band (4.500 to 6.000 cm^{-1}). In these regions of the spectrum and using the raw spectrum, the models misclassified two *M. aegyptia* stage 2 samples as *I. hederifolia* stage 2 (Table 3).

Table 3. PC-LDA matrix and hit percentage for species and stages in the spectrum ranges from 4.000 to 10.000 (full spectrum), 4.500 to 6.000 + 6.500 to 7.750, 4.500 to 6.000 and 6.500 to 7.750 cm^{-1} without pre-processing

Species and stages[1]	IH1 (n=21)	IH2 (n=31)	IH3 (n=35)	MA1 (n=30)	MA2 (n=22)	MA3 (n=41)	Hit (%)
All spectrum (4.000 a 10.000 cm^{-1})							
IH1	21	0	0	0	0	0	100.00
IH2	0	29	0	0	2	0	93.55
IH3	0	0	35	0	0	0	100.00
MA1	0	0	0	29	0	0	96.67
MA2	0	0	0	0	20	0	90.91
MA3	0	2	0	1	0	41	100.00
1st band + 2nd band (4.500 to 6.000 + 6.500 to 7.750 cm^{-1})							
IH1	21	0	0	0	0	0	100.00
IH2	0	31	0	0	2	0	100.00
IH3	0	0	35	0	0	0	100.00
MA1	0	0	0	30	0	0	100.00
MA2	0	0	0	0	20	0	90.91
MA3	0	0	0	0	0	41	100.00
1st band (4.500 a 6.000 cm^{-1})							
IH1	21	0	0	0	0	0	100.00
IH2	0	31	0	0	2	0	100.00
IH3	0	0	35	0	0	0	100.00
MA1	0	0	0	30	0	0	100.00
MA2	0	0	0	0	20	0	90.91
MA3	0	0	0	0	0	41	100.00
2nd band (6.500 a 7.750 cm^{-1})							
IH1	21	0	0	0	0	0	100.00
IH2	0	30	0	0	0	0	96.77
IH3	0	0	35	0	0	0	100.00
MA1	0	0	0	29	0	1	96.67
MA2	0	1	0	0	22	0	100.00
MA3	0	0	0	1	0	40	97.56

[1] Species: *I. hederifolia* in 10 stages (IH1); 11 (IH2) and 12 (IH3); and *M. aegyptia* in stages 11 (MA1); 12 (MA2) and 14 (MA3).

When applying PC-LDA in the ranges with SNV+1SG pre-processing (Table 4), the correct answers ranged from 40.91 to 96.77%. For the study of the bands, the results were superior, mainly for the 2nd band (6.500 to 7.750 cm^{-1}) which presented correct answers ranging from 90.91 to 100%.

Table 4. PC-LDA matrix and hit percentage for species and stages in the spectrum ranges from 4.000 to 10.000 (full spectrum), 4.500 to 6.000 + 6.500 to 7.750, 4.500 to 6.000 and 6.500 to 7.750 cm^{-1} with SNV+1SG pre-processing

Species and stages[1]	IH1 (n=21)	IH2 (n=31)	IH3 (n=35)	MA1 (n=30)	MA2 (n=22)	MA3 (n=41)	Hit (%)
All spectrum (4.000 a 10.000 cm^{-1})							
IH1	20	0	0	0	5	0	95.24
IH2	1	30	9	0	6	4	96.77
IH3	0	0	18	0	0	4	51.43
MA1	0	0	0	29	2	4	96.67
MA2	0	0	3	0	9	2	40.91
MA3	0	1	5	1	0	27	65.85
1st band + 2nd band (4.500 to 6.000 + 6.500 to 7.750 cm^{-1})							
IH1	21	0	0	0	0	0	100.00
IH2	0	31	0	0	5	0	100.00
IH3	0	0	34	0	0	0	97.14
MA1	0	0	1	30	0	0	100.00
MA2	0	0	0	0	17	0	77.27
MA3	0	0	0	0	0	41	100.00
1st band (4.500 a 6.000 cm^{-1})							
IH1	21	0	0	0	0	0	100.00
IH2	0	31	0	0	4	0	100.00
IH3	0	0	34	0	0	0	97.14
MA1	0	0	1	29	0	0	96.67
MA2	0	0	0	0	18	0	81.82
MA3	0	0	0	1	0	41	100.00
2nd band (6.500 a 7.750 cm^{-1})							
IH1	21	0	0	0	0	0	100.00
IH2	0	31	0	0	2	0	100.00
IH3	0	0	34	0	0	0	97.14
MA1	0	0	1	30	0	0	100.00
MA2	0	0	0	0	20	0	90.91
MA3	0	0	0	0	0	41	100.00

[1] Species*: I. hederifolia* in 10 stages (IH1); 11 (IH2) and 12 (IH3); and *M. aegyptia* in stages 11 (MA1); 12 (MA2) and 14 (MA3).

Although it is necessary to apply pre-processing, such as SNV+1SG, for unsupervised analysis (PCA), when applying supervised analysis (PC-LDA) the use of these pre-processing is unnecessary. Therefore, to discriminate *I. hederifolia* from *M. aegyptia* and its stages using PC-LDA analysis, it is recommended to study the 1st band (4.500 to 6.000 cm^{-1}) of the spectrum without pre-processing (Tables 3 and 4).

PCA grouped growth stages when applied to individual species and when applied to separate species into stages 1 and 2. PC-LDA analysis correctly classified species and stages with a percentage of success of up to 100%. The best discrimination results were observed in the 1st band (4.500 to 6.000 cm^{-1}) and 1st band + 2nd band (4.500 to 6.000 + 6.500 to 7.750 cm^{-1}) intervals (Table 3).

Discussion

Characterization of plant material and near infrared spectra

Although both species (*I. hederifolia* and *M. aegyptia*) were under the same management conditions, these variations were expected as a result of tissue turgor being characteristic of each species and stage (Barroso *et al.,* 2017).

The regions of the peaks in figure 5a coincide with the water absorbance peaks, which are close to 6.800 and 5.500 cm^{-1}, indicating the high moisture contents in the morning glory samples (Cunha Júnior *et al*., 2015; Torres *et al.*, 2019).

The water absorption range is quite characteristic in the infrared spectrum, as OH vibrations are easily detected (Pasquini, 2003; Wang *et al*., 2015). One of the first applications of diffuse reflectance spectroscopy in the infrared range was precisely with the purpose of exploring alternative methods for the determination of moisture in a non-destructive way (Pasquini, 2003; Torres *et al*., 2019).

However, although OH vibrations mostly explain the high absorbance peaks, in these ranges other organic compounds are also identified, such as flavonoids, alkaloids and anthocyanins (León and Downey, 2006; Jintao *et al*., 2018; Amanah *et al*., 2020). Pasquini (2003) emphasizes that the regions with the highest absorbance are those that contain the most considerable variations for understanding the spectral behavior of the samples. Therefore, in order to improve the understanding of the areas of greatest influence, the bands with the highest peaks and the highest values of loads were selected for the study of the multivariate analyses (Figure 6).

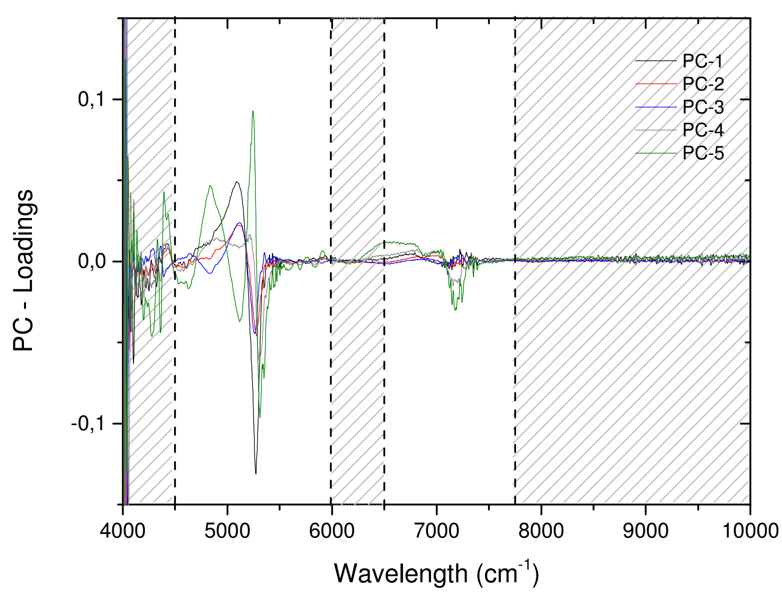

Figure 6. PC-loadings from 4.000 to 10.000 cm^{-1} with the pre-process

Principal Component Analysis (PCA)

The PCA results for each species reinforce that the spectral behavior of the species is distinct and that, although it is possible to study mixed models (with different groups of species), for PCA the best results are obtained when the analysis is applied to a single species. However, for practical application, the ideal would be to achieve high accuracies by applying mixed models, since, in the field, a single species will rarely be found in a single stage (Souza *et al*., 2020).

This result highlights that the differences in the chemical composition of the species are accentuated when the water content is reduced, contributing to a better grouping of samples (Figure 7).

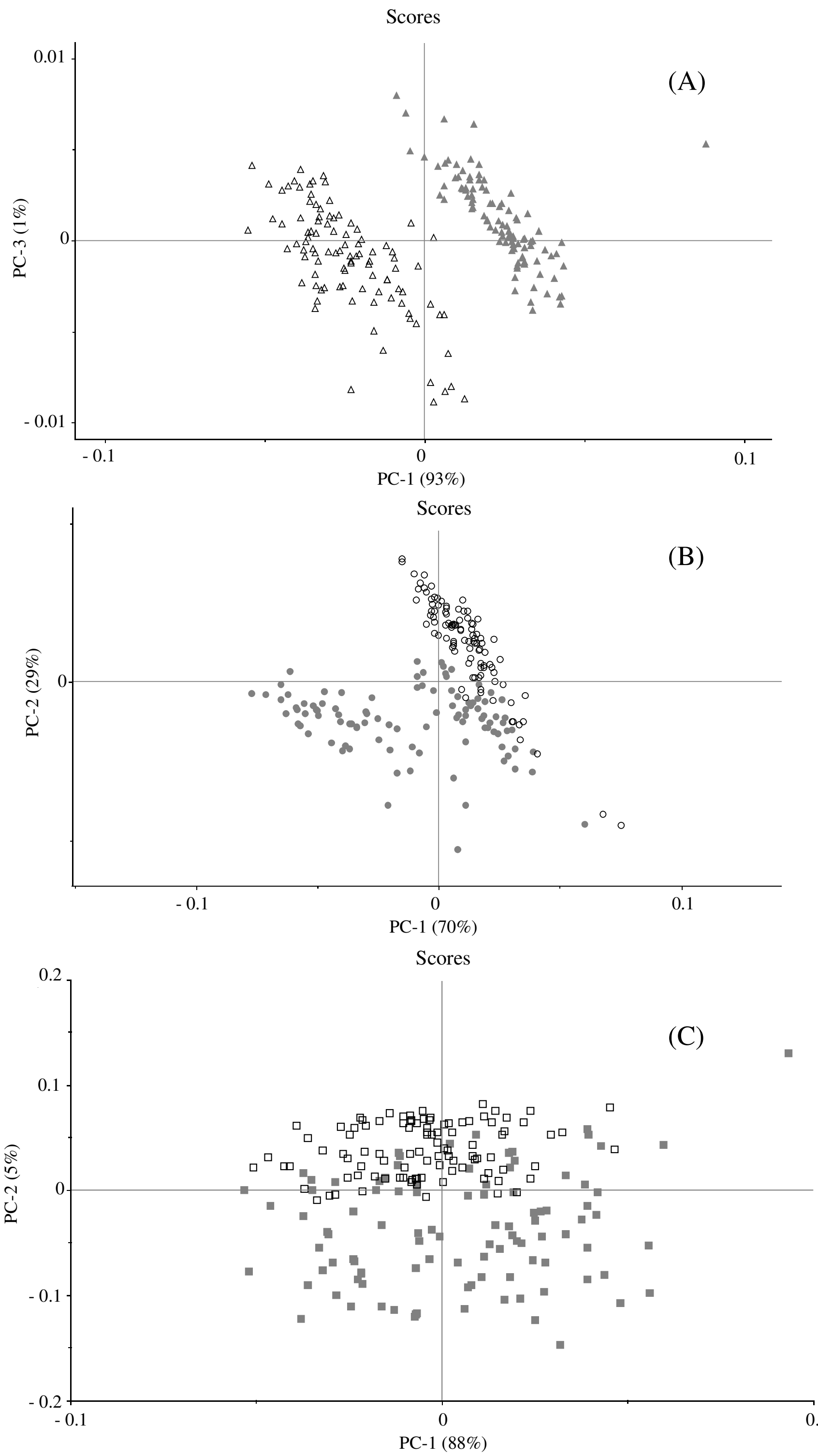

Figure 7. Values of the principal components resulting from the analysis of principal component (PCA) for the 1st spectrum band (4.500 to 6.000 cm^{-1}) for the species *I. hederifolia* (grey) and *M. aegyptia* (black) in the stages 1(a), 2(b) and 3(c)

The use of NIRS for monitoring and evaluating the quality of agricultural products is widely used in dry samples or with low moisture content, mainly to improve the understanding of the chemical composition and mitigate the interference of water in the spectrum (Cunha Júnior *et al*., 2015; Izadiyan *et al*., 2018; Giraudo *et al*., 2019). Therefore, for PCA applied to spectral data from fresh leaves, sample segregation is generally associated with water content (León and Downey, 2006; Torres *et al*., 2019).

LDA-based algorithms are very simple compared to other algorithms found in the literature. In addition, these algorithms significantly reduce data, speeding up computational analysis (Morais and Lima, 2018). Therefore, to achieve more promising results, the use of supervised analyses, such as PC-LDA, in this case, are quite common (Morais and Lima, 2018; Borraz-Martínez *et al*., 2019; Souza *et al*., 2020).

Linear discriminant analysis (PC-LDA)

The results presented by PC-LDA for species and stages also point to the high sensitivity that the technique has when combined with multivariate analyses (Table 3 and 4). Small differences in the stages of the species were enough to discriminate them with up to 100% precision. Thus, the next studies to identify weed species should pay attention to the construction of even more robust models, with even greater variations, since in this study the species were conducted under semi-controlled conditions, without considering the adversities that may be found in biotic and abiotic stress situations commonly observed in the field.

The results of this study demonstrated that NIR spectroscopy is able to discriminate species *Ipomoea hederifolia* and *Merremia aegyptia* at different growth stages with a remarkable accuracy, up to 100%. The best hit rates in PC-LDA were in the 1st range (4.500 to 6.000 cm^{-1}) and in the added ranges (4.500 to 6.000 + 6.500 to 7.750 cm^{-1}) for the spectrum without pre-processing (Table 3). As for the spectrum with SNV+1SG pre-processing, the best results were for the 2nd band (6.500 to 7.750 cm^{-1}) (Table 4).

Conclusions

The findings of this study demonstrate that NIR spectroscopy can effectively differentiate between the species *Ipomoea hederifolia* and *Merremia aegyptia* at various growth stages with remarkable accuracy, reaching up to 100%. This study represents an advance in the research and implementation of NIRS technology to discriminate weed species for the future development of equipment to assist in the adoption and/or performance of a specific management of weeds, capable of contributing to the reduction in the use of herbicides in crops.

Authors' Contributions

Conceptualization: AFB, LCCJ, GHAT and LCAA; Methodology: AFB, LCCJ, GHAT and PLCAA; Software: AFB and LCCJ; Validation: AFB and LCCJ; Formal analysis: AFB and LCCJ; Investigation: AFB; TGA and LCCJ; Data Curation: AFB, TGA and LCCJ; Resources: TGA and LCCJ; Supervision: LCCJ; Funding acquisition: LCCJ; Writing - Original Draft: AFB; TGA and LCCJ; Writing - Review & Editing: AFB; LCC; DPCS, LCCJ, GHAT and PLCAA; Project administration: AFB, LCCJ, GHAT and PLCAA.

All authors read and approved the final manuscript.

Ethical approval (for researches involving animals or humans)

Not applicable.

Acknowledgements

The authors would like to acknowledge the financial support of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (process no. 142375/2017-9).

Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

References

Alves FC, Bianco S, Nepomuceno MP, Martins JVF, Alves PLCA (2018). Shoot and root interference of morning glory on the initial growth of sugarcane. Anais da Academia Brasileira de Ciências 90:521-528. *https://doi.org/10.1590/0001-3765201820160233*

Amanah HZ, Joshi R, Masithoh RE, Choung MG, Kim KH, Kim G, Cho BK (2020). Nondestructive measurement of anthocyanin in intact soybean seed using Fourier Transform Near-Infrared (FT-NIR) and Fourier Transform Infrared (FT-IR) spectroscopy. Infrared Physics & Technology 111:103477. *https://doi.org/10.1016/j.infrared.2020.103477*

Barroso AAM, Ferreira PSH, Martins D (2019). Growth and development of *Ipomoea* weeds. Planta Daninha 37:1-9. *https://doi.org/10.1590/S0100-83582019370100034*

Bhullar MS, Walia US, Singh S, Singh M, Jhala AJ (2012). Control of Morningglories (*Ipomoea* spp.) in sugarcane (*Saccharum* spp.). Weed Technology 26:77-82. *https://doi.org/10.16l4/WT-D-l1-00112.1*

Bleiholder H, Kirfel H, Langelüddeke P, Stauss R (1991). Codificação Unificada dos Estádios Fenológicos de Culturas e Ervas Daninhas. Pesquisa Agropecuária Brasileira 26:1423-1429. *https://doi.org/10.1590/S1678-3921.pab1991.v26.3483*

Borraz-Martínez S, Simó J, Gras A, Mestre M, Boqué R (2019). Multivariate classification of *Prunus dulcis* varieties using leaves of nursery plants and near infrared spectroscopy. Scientific Reports 9:1-9. *https://doi.org/10.1038/s41598-019-56274-5*

Braga A, Chiconi L, Bacha A, Teixeira G, Cunha Junior L, Alves P (2023). Discrimination of morningglory species (*Ipomoea* spp.) using near-infrared spectroscopy and multivariate analysis. Weed Science 71(2):104-111. *https://doi.org/10.1017/wsc.2023.6*

Cunha Junior LC, Nardini V, Khatiwada BP, Teixeira GHA, Walsh KB (2015). Classification of intact açaí (*Euterpe oleracea* Mart.) and juçara (*Euterpe edulis* Mart) fruits based on dry matter content by means of near infrared spectroscopy. Food Control 50:630-636. *https://doi.org/10.1016/j.foodcont.2014.09.046*

Giraudo, A, Grassi S, Savorani F, Gavoci G, Casiraghi E, Geobaldo F (2019). Determination of the geographical origin of green coffee beans using NIR spectroscopy and multivariate data analysis. Food Control 99:137-145. *https://doi.org/10.1016/j.foodcont.2018.12.033*

Hansen KD, Garcia-Ruiz F, Kazmi W, Bisgaard M, La Cour-Harbo A, Rasmussen J, Andersen HJ (2013). An autonomous robotic system for mapping weeds in fields. IFAC Proceedings 46:217-224. *https://doi.org/10.3182/20130626-3-AU-2035.00055*

Hoagland D, Arnon D (1950). The water-culture method for growing plants without soil.: University of California, Berkeley.

Izadiyan P, Hemmateenejad B, Mohsen Taghavi S, Izadiyan M (2018). Discrimination of Shirazi thyme from thymus species and antioxidant activity prediction using chemometrics and FT-IR spectroscopy. Journal of the Iranian Chemical Society 15:259-268. *https://doi.org/10.1007/s13738-017-1228-4*

Jintao X, Yufei L, Liming Y, Chunyan L, Quanwei Y, Weiying W, Yun J, Minxiang Z, Peng L (2018). Rapid and simultaneous analysis of five alkaloids in four parts of Coptidis Rhizoma by near-infrared spectroscopy.

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 188:611-618. *https://doi.org/10.1016/j.saa.2017.07.053*

Klerkx L, Jakku E (2019). Labarthe P A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. NJAS - Wageningen Journal of Life Sciences 100315:90-91. *https://doi.org/10.1016/j.njas.2019.100315*

León L, Downey G (2006). Preliminary studies by visible and near-infrared reflectance spectroscopy of juvenile and adult olive (*Olea europaea* L.) leaves. Journal of the Science of Food and Agriculture 86:999-1004. *https://doi.org/10.1002/jsfa.2448*

Li H, Jiang D, Cao J, Zhang D (2020). Near-infrared spectroscopy coupled chemometric algorithms for rapid origin identification and lipid content detection of *Pinus koraiensis* seeds. Sensors 20:4905. *https://doi.org/10.3390/s20174905*

López-Granados F (2011). Weed detection for site-specific weed management: Mapping and real-time approaches. Weed Research 151:1-11. *https://doi.org/10.1111/j.1365-3180.2010.00829.x*

Morais CLM, Lima KMG (2018). Principal component analysis with linear and quadratic discriminant analysis for identification of cancer samples based on mass spectrometry. Journal of the Brazilian Chemical Society 29:472-481. *https://doi.org/10.21577/0103-5053.20170159*

Nieto-Ortega S, Lara R, Foti G, Melado-Herreros Á, Olabarrieta I (2023). Applications of near-infrared spectroscopy (NIRS) in fish value chain. In: El-Azazy M, Al-Saad K, El-Shafie AS (Eds). Infrared Spectroscopy - Perspectives and Applications. IntechOpen. *https://doi.org/10.5772/intechopen.100794*

Nugraha DT, Zaukuu JLZ, Bósquez JPA, Bodor Z, Vitalis F, Kovacs Z (2021). Near-infrared spectroscopy and aquaphotomics for monitoring mung bean (*Vigna radiata*) sprout growth and validation of ascorbic acid content. Sensors 21:611. *https://doi.org/10.3390/s21020611*.

Pagnoncelli FB, Trezzi MM, Brum B, Vidal RA, Portes AF, Scalcon EL, Machado A (2017). Morning glory species interference on the development and yield of soybeans. Bragantia 76:470-479. *https://doi.org/10.1590/1678-4499.2016.338*

Pasquini C (2003). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. Journal of the Brazilian Chemical Society 14:198-219. *https://doi.org/10.1590/S0103-50532003000200006*

Piza CST, Nepomuceno MP, Alves PLDCA (2016). Period prior to interference of morning glory in sugarcane. Científica 44:543. *https://doi.org/10.15361/1984-5529.2016v44n4p543-548*

Ribeiro NM, Torres BA, Ramos SK, Santos PHV, Simões CT, Monquero PA (2018). Differential susceptibility of morning glory (*Ipomoea* and *Merremia*) species to residual herbicides and the effect of drought periods on efficacy. Australian Journal of Crop Science 12:1090-1098. *https://doi.org/10.21475/ajcs.18.12.07.PNE1013*

Singh V, Rana A, Bishop M, Filippi AM, Cope D, Rajan N, Bagavathiannan M (2020). Unmanned aircraft systems for precision weed detection and management: Prospects and challenges. Advances in Agronomy 159:93-134. *https://doi.org/10.1016/bs.agron.2019.08.004*

Sohn SI, Pandian S, Oh YJ, Zaukuu JZ, Kang HJ, Ryu TH, Cho WS, Cho YS, Shin EK, Cho BK (2021). An overview of near infrared spectroscopy and its applications in the detection of genetically modified organisms. International Journal of Molecular Sciences 14:9940. *https://doi.org/10.3390/ijms22189940*

Souza MF, Amaral LR, Oliveira SRM, Coutinho MAN, Ferreira Netto C (2020). Spectral differentiation of sugarcane from weeds. Biosystems Engineering 190:41-46. *https://doi.org/10.1016/j.biosystemseng.2019.11.023*

Torres I, Sánchez MT, Benlloch-González M, Pérez-Marín D (2019). Irrigation decision support based on leaf relative water content determination in olive grove using near infrared spectroscopy. Biosystems Engineering 180:50-58. *https://doi.org/10.1016/j.biosystemseng.2019.01.016*

Van Evert FK, Fountas S, Jakovetic D, Crnojevic V, Travlos I, Kempenaar C (2017). Big Data for weed control and crop protection. Weed Research 57:218-233. *https://doi.org/10.1111/wre.12255*

Vitalis F, Zaukuu JLZ, Bodor Z, Aouadi B, Hitka G, Kaszab T, Kovacs Z (2020). Detection and quantification of tomato paste adulteration using conventional and rapid analytical methods. Sensors 20:6059. *https://doi.org/10.3390/s20216059*

Wang N, Zhang N, Wei J, Stoll Q, Peterson DE (2007). A real-time, embedded, weed-detection system for use in wheat fields. Biosystems Engineering 98:276-285. *https://doi.org/10.1016/j.biosystemseng.2007.08.007*

Wang P, Yu Z (2015). Species authentication and geographical origin discrimination of herbal medicines by near infrared spectroscopy: A review. Journal of Pharmaceutical Analysis 5:277-284. *https://doi.org/:10.1016/j.jpha.2015.04.001*