

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP  
CÂMPUS DE JABOTICABAL**

**ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO PARA  
DISCRIMINAÇÃO DE ESPÉCIES DE CORDAS-DE-VIOLA**

**Andreísa Flores Braga**  
**Engenheira Agrônoma**

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP  
CÂMPUS DE JABOTICABAL**

**ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO PARA  
DISCRIMINAÇÃO DE ESPÉCIES DE CORDAS-DE-VIOLA**

**Andreísa Flores Braga**

**Orientador: Prof. Dr. Pedro Luis da Costa Aguiar Alves**

**Coorientador: Prof. Dr. Luis Carlos da Cunha Júnior**

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutora em Agronomia (Produção Vegetal).

B813e Braga, Andreísa Flores  
Espectroscopia no infravermelho próximo para discriminação de espécies de cordas-de-viola / Andreísa Flores Braga. -- Jaboticabal, 2021  
70 p. : il., tabs., fotos

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal  
Orientador: Pedro Luis da Costa Aguiar Alves  
Coorientador: Luis Carlos da Cunha Júnior

1. Agronomia. 2. Ervas daninhas. 3. Convolvulaceae. 4. Quimiometria. 5. Agricultura de precisão. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

## CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO PARA DISCRIMINAÇÃO DE ESPÉCIES DE CORDAS-DE-VIOLA

AUTORA: ANDREÍSA FLORES BRAGA

ORIENTADOR: PEDRO LUÍS DA COSTA AGUIAR ALVES

COORDENADOR: LUIS CARLOS CUNHA JÚNIOR

Aprovada como parte das exigências para obtenção do Título de Doutora em AGRONOMIA (PRODUÇÃO VEGETAL), pela Comissão Examinadora:



Prof. Dr. PEDRO LUÍS DA COSTA AGUIAR ALVES (Participação Virtual)  
Departamento de Biologia Aplicada à Agropecuária / FCAV / UNESP - Jaboticabal

Prof. Dr. JOÃO MARTIM DE PORTUGAL E VASCONCELOS FERNANDES (Participação Virtual)  
Instituto Politécnico de Beja(ESA-IPBeja) / Beja/Portugal



Prof. Dr. GUSTAVO HENRIQUE DE ALMEIDA TEIXEIRA (Participação Virtual)  
Departamento de Ciências da Produção Agrícola (Produção Vegetal) / FCAV / UNESP - Jaboticabal



Prof. Dr. GILMARCOS DE CARVALHO CORRÊA (Participação Virtual)  
Universidade Federal de Goiás/UFG / Goiânia/GO



Prof. Dr. ARTHUR ARROBAS MARTINS BARROSO (Participação Virtual)  
Universidade Federal do Paraná-UFPR / Curitiba/PR



Jaboticabal, 15 de abril de 2021

## **DADOS CURRICULARES DA AUTORA**

**ANDREÍSA FLORES BRAGA** – nascida em 20 de dezembro de 1991, no município de Bom Sucesso, estado de Minas Gerais, filha de Vera Alice Aparecida Flores Braga e Manoel Geraldo Resende Braga. Fez sua graduação em Engenharia Agrônoma na Universidade Federal de Lavras (UFLA), com conclusão em julho de 2014. Durante a graduação, foi bolsista de iniciação científica da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), desenvolveu pesquisas relacionadas à produção de metabólitos secundários em plantas medicinais, no período de 2011 a 2014. Em agosto de 2014, iniciou o curso de mestrado em Agronomia (Produção Vegetal), na Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) – Campus de Jaboticabal. Durante o curso de mestrado, foi bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Conduziu sua pesquisa junto ao grupo de pesquisa LAPDA, e também contribuiu na elaboração e condução de projetos de pesquisa em parcerias com empresas e na organização de eventos para popularização e divulgação científica. No período de setembro a dezembro de 2015, desenvolveu parte de suas pesquisas na Universidad de la República (UDELAR) – Campus de Paysandú, Uruguai. Obteve o título de mestre em agosto de 2016, com a dissertação intitulada: “Desenvolvimento inicial do eucalipto em convivência com densidades de azevém resistente e suscetível ao glyphosate”. Após o mestrado, iniciou juntamente com a direção do campus e outros dois pós-graduandos, o projeto de implantação da incubadora de empresas de base tecnológica – Inovajab, realizando atividades de ensino e apoio à transferência de tecnologias e fortalecimento do ecossistema de inovação de Jaboticabal – SP, que culminou no reconhecimento da cidade como arranjo produtivo local do agronegócio, em maio de 2020. Iniciou o curso de doutorado em Agronomia (Produção Vegetal), na UNESP – Campus de Jaboticabal em agosto de 2017, desenvolvendo sua pesquisa junto ao grupo de pesquisa LAPDA – Laboratório de Plantas Daninhas e ao grupo de pesquisa LTTC – Laboratório de Tecnologia em Pós-colheita de Frutas e Hortaliças, da Universidade Federal de Goiás (UFG). Durante o curso de doutorado, foi bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

“E disse Deus: Haja luz; e houve luz.”

*Gênesis 1:3*

“O conhecido é finito, o desconhecido, infinito, intelectualmente estamos numa ilha no meio de um oceano ilimitado de inexplicabilidade. Nossa função em cada geração é reivindicar um pouco mais de terra firme.”

*Thomas Henry Huxley*

Aos meus pais, Manoel Geraldo e Vera Alice, e  
aos meus irmãos, Gabriel, Manoel Victor e André.

**Dedico.**

## **AGRADECIMENTOS**

A Deus, pela saúde, força, inspiração e criatividade.

Ao meu orientador, Prof. Dr. Pedro Luis da Costa Aguiar Alves, pela amizade, orientação, confiança e liberdade concedida durante minha trajetória na academia.

Ao meu coorientador, Prof. Dr. Luis Carlos da Cunha Júnior, pela orientação e sobretudo pela paciência ao ensinar e pelas palavras de estímulo ao longo de todo desenvolvimento do trabalho.

Ao Prof. Dr. Gustavo Henrique de Almeida Teixeira, pelo empréstimo do equipamento e pelas considerações nas bancas de qualificação e defesa.

Aos Dr. Marcos Antônio Kuva, Dr. João Martim de Portugal e Vasconcelos Fernandes, Dr. Gilmarcos de Carvalho Corrêa e Dr. Arthur Arrobas Martins Barroso, pelas considerações e participação nas bancas de qualificação e defesa.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, pelo auxílio financeiro para desenvolvimento desta pesquisa, número do processo: 142375/2017-9.

A todos meus professores, que fizeram parte da minha trajetória no curso de doutorado; compartilhar o conhecimento representa para mim a missão mais ilustre.

A todos os colaboradores da Universidade Estadual Paulista “Júlio de Mesquita Filho”, UNESP/FCAV - Campus de Jaboticabal, pela contribuição à minha formação profissional.

Aos coordenadores e membros do conselho do Programa de Pós-graduação em Agronomia (Produção Vegetal).

Aos funcionários da Seção Técnica de Pós-graduação, em especial ao Diego Mafra, pela atenção e solicitude.

Aos funcionários da biblioteca por todo suporte concedido, em especial a Luciane Meire Ribeiro e a Marli Saes.

A todos os membros e ex-membros do grupo de pesquisa LAPDA, por todo auxílio e convívio durante minha permanência no grupo; em especial ao Leandro Chiconi, pelo apoio durante a execução dos experimentos.

À Dr<sup>a</sup> Mariluce Pascoina Nepomuceno, pelas orientações, amizade e longas conversas sobre a ciência e a vida, que me inspiraram e me fizeram sentir acolhida.

Aos servidores do Departamento de Biologia Aplicada à Agricultura, pela atenção e apoio durante todo o curso.

Ao servidor técnico e amigo José Valcir Fidelis Martins, pela valiosa ajuda nos experimentos e pelas longas conversas e momentos de descontração.

Aos colegas da Inovajab, Diego Siqueira, Rafael Parras e Ingrid Martins, pelo convívio e estímulo ao meu desenvolvimento profissional e pessoal.

A todos meus amigos, em especial aos meus companheiros de jornada acadêmica Allan Bacha, Ana Stella Gonçalves, Andreísa Lima, Bruna Pires, Cristiane Pereira, Isa Marcela Braga, Izabela Orzari, Juciléia Wagatsuma, Juliana Rodrigues, Neriane Hijano, Pedro Martins, Renata Thaysa Santos, Thiago Oliveira e Willians Carrega.

Ao Marcelo Campos, pelo companheirismo, apoio e disposição em sempre me ajudar.

Aos meus pais e irmãos por sempre acreditarem nas minhas escolhas e me apoiarem.

Por fim, a todos aqueles que passaram pela minha vida nesses últimos anos e fizeram parte da minha trajetória, contribuindo de alguma forma para a realização deste trabalho.

Muito obrigada!



|                                                                                                                                                        |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.4. Seleção de faixas espectrais.....                                                                                                                 | 41 |
| 4. CONCLUSÃO.....                                                                                                                                      | 44 |
| 5. REFERÊNCIAS.....                                                                                                                                    | 45 |
| CAPÍTULO 3 – Espectroscopia NIR e métodos multivariados para discriminação de espécies de cordas-de-violão em diferentes estádios de crescimento ..... | 47 |
| 1. INTRODUÇÃO .....                                                                                                                                    | 48 |
| 2. MATERIAL E MÉTODOS .....                                                                                                                            | 49 |
| 2.1. Material Vegetal.....                                                                                                                             | 49 |
| 2.2. Obtenção dos espectros.....                                                                                                                       | 51 |
| 2.3. Quimiometria .....                                                                                                                                | 52 |
| 3. RESULTADOS E DISCUSSÃO.....                                                                                                                         | 52 |
| 3.1. Caracterização do material vegetal e espectros NIR.....                                                                                           | 52 |
| 3.2. Análise de componente principal (PCA) .....                                                                                                       | 55 |
| 3.3. Análise linear discriminante (PC-LDA) .....                                                                                                       | 60 |
| 4. CONCLUSÃO.....                                                                                                                                      | 64 |
| 5. REFERÊNCIAS.....                                                                                                                                    | 66 |
| CAPÍTULO 4 - Considerações finais .....                                                                                                                | 69 |

## ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO PARA DISCRIMINAÇÃO DE ESPÉCIES DE CORDAS-DE-VIOLA

**RESUMO** - A ocorrência de plantas daninhas em áreas agrícolas é um dos principais fatores que limitam a produtividade e aumentam os custos de produção. O manejo específico destas espécies tem ganhado destaque nos últimos anos, principalmente pelo avanço das tecnologias e por apresentar uma alternativa à redução do uso de herbicidas e exposição ambiental. O grande desafio para o manejo específico de plantas daninhas é identificar as espécies alvo em meio às plantas cultivadas. Neste contexto, diversas tecnologias têm sido estudadas para identificação e reconhecimento destas espécies em campo. A espectroscopia no infravermelho próximo (NIRS) é bastante empregada em análises de qualidade e caracterização de alimentos e solos, mas apesar de ser uma técnica bastante difundida, ainda são poucos os estudos relacionados ao seu uso para identificação/discriminação de plantas daninhas. O objetivo deste trabalho foi verificar o potencial discriminatório da espectroscopia NIR associado a análises multivariadas em espécies de cordas-de-viola e estádios de crescimento. Para isso, realizou-se experimentos com os seguintes objetivos: (I) discriminar as espécies de *Ipomoea hederifolia*, *Ipomoea nil* e *Merremia aegyptia*, (II) discriminar as espécies de *I. hederifolia* e *M. aegyptia*, considerando três estádios de crescimento. A semeadura das espécies foi realizada em vasos plásticos com capacidade para 0.5 litros, preenchidos com areia de rio lavada e peneirada. As espécies foram cultivadas em ambiente aberto e irrigadas diariamente com água e solução nutritiva. No dia da coleta dos espectros, as plantas foram caracterizadas quanto ao teor de água e ao estágio de crescimento. Os espectros NIRS foram adquiridos em absorbância na faixa espectral de 10.000 a 4.000  $\text{cm}^{-1}$  (1.000 a 2.500 nm). No momento da coleta dos espectros NIR, a temperatura do ambiente foi mantida a  $23\text{ }^{\circ}\text{C} \pm 1$  e a leitura dos espectros foi realizada na parte adaxial das folhas frescas, posicionadas sobre um fundo preto. Os dados espectrais foram analisados usando o programa Unscrambler®. Pré-processamentos e faixas espectrais foram testados e os dados espectrais foram submetidos as análises de componente principal (PCA); discriminante linear de componentes principais (PC-LDA) e apenas no primeiro experimento, a análise discriminante por mínimos quadrados parciais (PLS-DA). O estudo das faixas espectrais bem como os pré-processamentos foram relevantes e promoveram melhorias na capacidade de acerto dos modelos. Os valores de acurácia no primeiro experimento foram de até 99,26, 98,52 e 98,73%, para *I. hederifolia*, *I. nil* e *M. aegyptia*, respectivamente. No segundo experimento as melhores faixas para discriminar as espécies e seus respectivos estádios foram de 4.500 a 6.000  $\text{cm}^{-1}$  e de 4.500 a 6.000 + 6.500 a 7.750  $\text{cm}^{-1}$ , com assertividade de até 100% para maioria das classes, sem a necessidade de pré-processamento. Os resultados destes estudos demonstraram que a espectroscopia NIR pode ser utilizada para discriminar espécies de corda-de-viola em diferentes estádios de crescimento, representando um avanço na pesquisa e implantação desta técnica para o desenvolvimento de equipamentos que auxiliem a adoção e/ou realização de um manejo específico de plantas daninhas.

**Palavras-chave:** agricultura de precisão, assinatura espectral, identificação de plantas daninhas, manejo específico, quimiometria

## NEAR INFRARED FOR DISCRIMINATION OF MORNING GLORY SPECIES

**ABSTRACT** – The occurrence of weeds in cultivated areas is one of the main factors that limit agricultural productivity and increase production costs. The specific management of these species has gained prominence in recent years, mainly due to the advancement of technologies and for presenting an alternative to reduce the use of herbicides and environmental exposure. The great challenge for specific weed management is to identify the target species among the cultivated plants. In this context, several technologies have been studied for species identification and field recognition. Near-infrared spectroscopy (NIRS) is widely used in the analysis of quality and characterization of food and soils, but despite being a widespread technique, there are still few studies related to its use for weed identification/discrimination. The objective of this work was to verify the discriminatory potential of NIRS associated with multivariate analyzes in species of morning glory and growth stages. Thus, experiments were carried out with the following objectives: (I) discriminating the species of *Ipomoea hederifolia*, *Ipomoea nil*, and *Merremia aegyptia*, using multivariate techniques and pre-processing and spectral range selection; (II) discriminate the species of *I. hederifolia* and *M. aegyptia*, considering three growth stages. The sowing of the species was carried out in plastic pots with a capacity of 0.5 L, filled with sieved river sand, previously washed. The species were grown in an open environment and irrigated daily with water and nutrient solution. On the day of spectra collection, the plants were characterized for water content and stage (in the second experiment). The NIR spectra were acquired in absorbance in the spectral range of 10,000 to 4,000  $\text{cm}^{-1}$  (1,000 to 2,500 nm). At the time of collection of the NIR spectra, the ambient temperature was maintained at  $23^{\circ}\text{C} \pm 1$  and the reading of the spectra was performed on the adaxial part of the fresh leaves, positioned on a black background. The spectral data were analyzed using the Unscrambler<sup>®</sup> program. The pre-processing and spectral bands were tested and the spectral data submitted to the principal component analysis (PCA); linear discriminant of principal components (PC-LDA) and only in the first experiment, the discriminant analysis of partial least squares (PLS-DA). The study of the spectral ranges as well as the pre-processing were relevant and promoted improvements in the accuracy of the models. The accuracy values in the first experiment were up to 99.26, 98.52, and 98.73%, for *I. hederifolia*, *I. nil* and *M. aegyptia*, respectively. In the second experiment, the best intervals to discriminate species and their respective stages were 4,500 to 6,000  $\text{cm}^{-1}$  and 4,500 to 6,000 + 6,500 to 7,750  $\text{cm}^{-1}$ , with assertiveness of up to 100% for most classes, without the need for pre-processing. The results of these studies demonstrated that NIRS spectroscopy can discriminate species of morning glory at different stages of growth, representing an advance in the research and implementation of this technique for the development of equipment that assists in the adoption of specific management.

**Keywords:** chemometrics, precision agriculture, specific management, spectral signature, weed identification

## LISTA DE ABREVIATURAS

ANOVA – análise de variância

BBCH – sigla em alemão de “Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie”, escala usada para identificar os estágios de desenvolvimento fenológico das plantas

DT – Detrend (tratamento matemático)

EVA - Espuma de vinil

FN – falso negativo

FP – falso positivo

IH – *Ipomoea hederifolia*

IH1 - *Ipomoea hederifolia* no microestádio 10, conforme o BBCH

IH2 - *Ipomoea hederifolia* no microestádio 11, conforme o BBCH

IH3 – *Ipomoea hederifolia* no microestádio 12, conforme o BBCH

IN – *Ipomoea nil*

LDA - análise linear discriminante

MA – *Merremia aegyptia*

MA1 - *Merremia aegyptia* no microestádio 11, conforme o BBCH

MA2 - *Merremia aegyptia* no microestádio 12, conforme o BBCH

MA3 - *Merremia aegyptia* no microestádio 14, conforme o BBCH

NIR - infravermelho próximo

NIRS - espectroscopia no infravermelho próximo

PC – componente principal

PCA - análise de componente principal

PC-LDA - análise linear discriminante do componente principal

PLS-DA - análise discriminante por quadrados mínimos parciais

SEN – sensibilidade

SG - Savitzky-Golay (tratamento matemático)

SNV - variação normal padrão

SPEC – especificidade

TA – teor de água

TEF - taxa de eficiência

TFN - taxa de falsos negativos

TFP - taxa de falsos positivos

VN –verdadeiro negativo

VP –verdadeiro positivo

## LISTA DE TABELAS

### CAPÍTULO 1

**Tabela 1.** Tabela de contingência ou matriz de confusão dos resultados de uma classificação. Fonte: Adaptado, Ferreira (2015).

### CAPÍTULO 2

**Tabela 1.** Resultados das análises discriminantes de PC-LDA e PLS-DA para as três espécies de corda-de-viola sem e com diferentes pré-processamentos na faixa espectral de 4.000-10.000  $\text{cm}^{-1}$ .

**Tabela 2.** Resultados da análise de PLS-DA para três espécies de corda-de-viola sem e com diferentes pré-processamento em diferentes faixas espectrais.

### CAPÍTULO 3

**Tabela 1.** Código decimal para a classificação dos estádios de desenvolvimento de plantas

**Tabela 2.** Valores médios do teor de água para *I. hederifolia* e *M. aegyptia* nos três estádios de crescimento.

**Tabela 3.** Matriz da PC-LDA e porcentagem de acerto para espécies e estádios nas faixas do espectro de 4.000 a 10.000 (espectro todo), 4.500 a 6.000 + 6.500 a 7.750, 4.500 a 6.000 e 6.500 a 7.750  $\text{cm}^{-1}$  sem pré-processamento.

**Tabela 4.** Matriz da PC-LDA e porcentagem de acerto para espécies e estádios nas faixas do espectro de 4.000 a 10.000 (espectro todo), 4.500 a 6.000 + 6.500 a 7.750, 4.500 a 6.000 e 6.500 a 7.750  $\text{cm}^{-1}$  com pré-processamento SNV+1SG.

## LISTA DE FIGURAS

### CAPÍTULO 1

**Figura 1.** Formas de interação da radiação eletromagnética com a superfície amostrada: (a) refletância; (b) transmitância; (c) interactância e (d) transfectância. Fonte: Elaborado pela autora, baseado em Pasquini, (2003).

**Figura 2.** Grupos funcionais atribuídos aos comprimentos de onda no infravermelho próximo. Fonte: (Adaptado, Xiabo et al., 2010).

**Figura 3.** Curva de refletância espectral típica de vegetação saudável representando diferentes picos de absorção. Fonte: Adaptado, Prabhakar et al., (2012).

**Figura 4.** Representação gráfica de dois grupos de amostras no espaço bidimensional definido pelas Variáveis 1 e 2 e as respectivas componentes principais (a); Representação gráfica das amostras projetadas em PC1 e PC2 (b). Fonte: Adaptado, Ferreira (2015).

### CAPÍTULO 2

**Figura 1.** Dados de absorbância do espectro todo, sem pré-processamento (a) e com pré-processamento SNV (b) de *I.hederifolia* (IH), *I.nil* (IN) e *M.aegyptia* (MA).

**Figura 2.** Valores de PC 1 e 2 para a análise de componente principal (PCA) para todo o espectro sem pré-processamento (a) e para a faixa de 4240-5540  $\text{cm}^{-1}$  com o pré-processamento SNV (b), para as espécies de *I.hederifolia* (quadrado), *I.nil* (triângulo) e *M.aegyptia* (círculo).

## CAPÍTULO 3

**Figura 1.** *Ipomoea hederifolia* (a) e *Merremia aegyptia* (b) nos estádios de crescimento 1, 2 e 3.

**Figura 2.** Absorbância média dos espectros (a) e PC-loadings (b) de 4.000 a 10.000  $\text{cm}^{-1}$  com o pré-processamento SNV+1SG das espécies de *I. hederifolia* nos estádios 10 (IH1); 11 (IH2) e 12 (IH3); e *M. aegyptia* nos estádios 11 (MA1); 12 (MA2) e 14 (MA3). (1ª faixa: 4.500 a 6.000  $\text{cm}^{-1}$ ; 2ª faixa: 6.500 a 7.750  $\text{cm}^{-1}$ ).

**Figura 3.** Valores de PC 1 e 2 resultantes da análise de componente principal (PCA) com o espectro todo (a) e com as duas faixas (b) para as espécies *I. hederifolia* (cinza) e *M. aegyptia* (preto) em três estádios de desenvolvimento. (triângulo: estágio 1; círculo: estágio 2; quadrado: estágio 3).

**Figura 4.** Valores de PC 1 e 2 resultantes da análise de componente principal (PCA) com a 1ª faixa (a) e com a 2ª faixa (b) para as espécies de *I. hederifolia* (cinza) e *M. aegyptia* (preto) em três estádios de desenvolvimento. (triângulo: estágio 1; círculo: estágio 2; quadrado: estágio 3).

**Figura 5.** Valores de PC 1 e 2 resultantes da análise de componente principal (PCA) na 1ª faixa do espectro (4500 a 6000  $\text{cm}^{-1}$ ) para as espécies de *I. hederifolia* (a) e *M. aegyptia* (b) nos três estádios de crescimento. (triângulo: estágio 1; círculo: estágio 2; quadrado: estágio 3).

**Figura 6.** Valores dos componentes principais resultantes da análise de componente principal (PCA) para a 1ª faixa do espectro (4500 a 6000  $\text{cm}^{-1}$ ) com as espécies *del.hederifolia* (cinza) e *M. aegyptia* (preto) nos estádios 1(a), 2(b) e 3(c)

## **CAPÍTULO 1 – Considerações gerais**

### **1. INTRODUÇÃO**

O Brasil desempenha papel fundamental no contexto de produção de alimentos no mundo. Por outro lado, também é desafiado a produzir cada vez mais de maneira sustentável, gerando menos impacto ao meio ambiente (Martinelli et al., 2010). Os desafios para o desenvolvimento de uma agricultura mais produtiva e sustentável são os agentes propulsores do desenvolvimento de tecnologia e inovação no setor.

O manejo de plantas daninhas nas lavouras é indispensável para o alcance de altas produtividades e a não realização de um manejo adequado acarreta perdas ainda maiores (Kuva et al., 2007; Santos et al., 2017). Nas últimas décadas, o principal manejo adotado para o controle de plantas daninhas tem sido o controle químico, realizado pela aplicação de herbicidas, geralmente aplicados em área total (Correia, 2016). Contudo, com a introdução de conceitos de agricultura de precisão e o desenvolvimento de novas tecnologias, o manejo destas espécies tem passado por algumas mudanças de paradigmas, envolvendo o uso mais racional e eficiente dos herbicidas (Pannacci et al., 2017).

O desenvolvimento de tecnologias nas áreas de computação e automação tem estimulado o surgimento de serviços e produtos baseados no uso de sensores, visão de máquina e uso de informações georreferenciadas capazes de distinguir plantas daninhas (Fernández-Quintanilla et al., 2018). Na etapa de desenvolvimento de novos protocolos de manejo é fundamental a integração de tecnologias antigas às novas, e que considerem principalmente as características biológicas e ecológicas dessas espécies (Klerkx et al., 2019; Van Evert et al., 2017).

A discriminação de plantas daninhas por meio de dados espectrais ou imagens é uma abordagem que pode auxiliar no manejo específico destas espécies, gerando economia no uso de herbicidas e menor exposição de moléculas ao ambiente (Souza et al., 2020). Alguns sensores de plantas daninhas já estão sendo usados comercialmente e são capazes de detectar a presença ou ausência de plantas sobre o solo. Porém, o desenvolvimento das tecnologias futuras está buscando a

diferenciação e o reconhecimento de espécies de plantas daninhas em meio as culturas (Fernández-Quintanilla et al., 2018).

As plantas daninhas dos gêneros *Ipomoea* e *Merremia* são espécies anuais, herbáceas e trepadeiras, que ganharam expressividade nos últimos anos, devido à implementação do sistema de colheita mecanizada em cana-de-açúcar e adoção de plantio direto em culturas anuais (Monquero et al., 2011; Correia, 2016). Essas espécies, além de competirem com as culturas por água, luz e nutrientes, interferem significativamente nas práticas culturais, especialmente na colheita, reduzindo a eficiência e impactando nos custos (Piza et al., 2016; Pagnoncelli et al., 2017). Apesar de serem espécies pertencentes à mesma família, por possuírem semelhanças morfológicas, elas respondem de modo diferenciado ao controle químico, ressaltando, assim, a importância de diferenciá-las no momento do controle.

A espectroscopia no infravermelho próximo (NIRS) associada a técnicas de quimiometria é aplicada na agricultura já há muito tempo (Pasquini, 2003; Cao e Cao, 2020). As principais aplicações envolvem análises de qualidade e caracterização de alimentos e solos (Marques et al., 2014; Cunha Júnior et al., 2015; Roza-Delgado et al., 2017; dos Santos Neto et al., 2017). Recentemente, estudos têm explorado seu potencial de discriminação em espécies de plantas, para fins de identificação de cultivares, rastreamento geográfico e discriminação de plantas daninhas (Soares et al., 2017; Steidle Neto et al., 2018; Snel et al., 2018; Souza et al., 2020).

Contudo, o presente trabalho apresenta métodos para análises espectrais e discriminação de espécies e estádios de corda-de-viola utilizando a espectroscopia NIR. O primeiro capítulo contextualiza a importância das espécies estudadas, a evolução de técnicas de controle e a aplicação de NIRS na agricultura. No segundo capítulo é abordado a discriminação de três espécies de corda-de-viola, sendo elas: *Ipomoea hederifolia*, *Ipomoea nil* e *Merremia aegyptia*. No terceiro capítulo é abordado a discriminação de *I. hederifolia* e *M. aegyptia*, considerando diferentes estádios de crescimento. Por fim, nas considerações finais, são colocados os avanços e perspectivas que este estudo traz dentro desta área do conhecimento.

**Hipótese**

A espectroscopia NIR é capaz de discriminar espécies e estádios de crescimento de plantas daninhas de diferentes gêneros e de uma mesma família.

**Objetivo**

Verificar o potencial discriminatório do uso de espectroscopia no infravermelho próximo associado a técnicas multivariadas em espécies de cordas-de-viola e estádios de crescimento distintos.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1. Família das Convolvulaceae

Convolvulaceae é uma família de espécies de plantas de grande distribuição, caracterizada pela alta diversidade de características morfológicas e habitats ecológicos (Stefanovic et al., 2003). Esta família é praticamente cosmopolita em distribuição, ocorrendo, sobretudo, em regiões tropicais e subtropicais, e engloba aproximadamente 1.600 - 1.800 espécies, agrupadas em 55 - 60 gêneros (Austin e Huaman, 1996).

As espécies de Convolvulaceae podem ser descritas como ervas trepadeiras, volúveis, lianas, arbustos ou árvores, com folhas frequentemente simples, pinadamente lobadas ou pectinadas, palmadamente compostas em algumas espécies ou reduzidas a escamas e, suas flores, podem ser hermafroditas ou unissexuadas, com botões frequentemente induplicados (Austin e Cavalcanti, 1982).

O gênero *Ipomoea*, contém entre 600-700 espécies, compreende o maior número de espécies dentro das Convolvulaceae, sendo que mais da metade destas ocorrem nas Américas (Austin e Huáman, 1996). Já o gênero *Merremia*, com aproximadamente 60 espécies, está presente em regiões tropicais e subtropicais de ambos os hemisférios (Austin e Cavalcanti, 1982).

As principais espécies de interesse econômico das Convolvulaceae, são a batata-doce (*Ipomoea batatas*) e o espinafre-d'água (*Ipomoea aquatica*), além das espécies com valor hortícola, porém menos expressivas, usadas para fins medicinais, terapêuticos e ornamentais (Lorenzi, 1991; Lorenzi e Matos, 2002; Felger et al, 2012).

Embora existam poucas espécies de interesse econômico na agricultura, as espécies dos gêneros *Ipomoea* e *Merremia*, são popularmente conhecidas como corda-de-viola e são bastante estudadas, devido à alta ocorrência em áreas agrícolas (Correia e Kronka, 2010; Monquero et al., 2011; Ribeiro et al., 2018). Entre as espécies mais ocorrentes, destacam-se: *I. hederifolia*, *I. quamoclit*, *I. nil*, *I. grandifolia*, *M. cissoides* e *M. aegyptia* (Azania et al., 2002; Kuva et al., 2007; Silva et al., 2009).

A importância destas plantas daninhas em áreas agrícolas aumentou nos últimos anos, principalmente devido às mudanças nos sistemas de manejo das

culturas, como por exemplo, a adoção do plantio direto em culturas anuais e da colheita mecanizada na cana-de-açúcar (Kuva et al., 2007; Ribeiro et al., 2018; Barroso et al., 2017).

O aumento da infestação de espécies de *Ipomoea* e *Merremia* em cana-de-açúcar tem se tornado crítico, principalmente por considerar que a manutenção da palhada na superfície do solo é capaz de fornecer a estas espécies um ambiente favorável para a germinação e desenvolvimento, devido às menores faixas de temperatura e maior conservação da umidade do solo (Correia, 2016). Além disso, espécies de corda-de-viola, até mesmo dentro de um mesmo gênero, possuem características biológicas distintas que afetam o grau de interferência às culturas e à suscetibilidade aos herbicidas (Orzari et al., 2013). Portanto, conhecer estas espécies, as etapas fenológicas, os hábitos de crescimento e desenvolvimento, o ciclo reprodutivo, bem como os fatores ecológicos, permitem a adoção de um manejo integrado e inferência da capacidade competitiva dessas espécies (Orzari et al., 2013; Barroso et al., 2019).

Os prejuízos ocasionados por essas espécies podem ser diretos e/ou indiretos. Os diretos estão relacionados à matocompetição, enquanto os indiretos estão relacionados às práticas culturais, principalmente a colheita mecanizada (Azania et al., 2002; Correia, 2016). Em cultivos de cana-de-açúcar, áreas com infestação predominante de *I. hederifolia* reduziram a produtividade dos colmos em até 46% (Silva et al., 2009). Já na cultura da soja, a presença de *I. grandifolia* e *I. purpurea*, chegaram a reduzir em até 80% a produtividade (Pagnoncelli et al., 2017).

O hábito de crescimento trepador faz com que essas espécies sejam ainda mais problemáticas, pois se entrelaçam nas folhas e colmos, afetando a eficiência fotossintética e impedindo a colheita mecanizada, acarretando o embuchamento da colhedora (Christoffoleti et al., 2006). Além disso, espécies de corda-de-viola possuem sementes grandes, com reservas, as quais podem ser arrastadas pelas colhedoras e germinar após grandes intervalos de tempo, mesmo sob grande quantidade de palhada (Kissmann e Groth, 1999; Silva et al., 2009; Correia, 2016).

A interferência das plantas daninhas varia em função das espécies, mas as plantas dos gêneros *Ipomoea* e *Merremia* são semelhantes quanto ao hábito de crescimento e por isso são predominantemente conhecidas pelo mesmo nome

popular de corda-de-viola, o que pode dificultar a correta identificação das espécies e consequentemente prejudicar na escolha adequada do herbicida (Campos et al., 2009). Portanto, compreender a biologia das plantas daninhas é fundamental para o desenvolvimento de técnicas adequadas que permitam a adoção de manejos eficazes.

## **2.2. Manejo de plantas daninhas**

A interferência de plantas daninhas nos cultivos agrícolas é uma das principais limitações enfrentadas pelos produtores, uma vez que afetam direta e indiretamente a produtividade (Azania et al, 2002; Kuva et al., 2007). Sendo assim, o alcance de altas produtividades está condicionado ao manejo eficaz dessas espécies, que pode ser realizado por métodos biológicos, culturais, mecânicos ou químicos, adotados de forma individual ou combinados (Pitelli, 1985). O ideal para o manejo de plantas daninhas é a adoção de um manejo integrado, que se baseia na integração dos métodos de controle, considerando todos os recursos disponíveis para reduzir a interferência pelas plantas daninhas (Owen et al., 2014; Pannacci et al., 2017).

A efetividade do controle de plantas daninhas varia em função de diversos fatores, como, por exemplo, a espécie infestante, a densidade de infestação, o estágio de desenvolvimento, a época de interferência, o sistema de cultivo, a cultura, o espaçamento e o tipo de solo (Piza et al., 2016; Pannacci et al., 2017; Braga et al., 2018). Em função do baixo custo e alta eficiência, o controle químico tem sido o mais adotado nos últimos anos para o controle de espécies de plantas daninhas (Owen et al., 2014; Westwood et al., 2018).

A adoção do controle químico consiste na aplicação de produtos chamados de herbicidas, capazes de interferirem nos processos bioquímicos e fisiológicos das plantas, matando ou retardando o crescimento destas (Pitelli, 1987). No entanto, embora seja uma das práticas mais economicamente viáveis, o uso destes produtos corresponde a mais de um terço do custo total de insumos (Oerke e Dehne, 2004; Santos et al, 2017). Isso porque, além de serem usados como dessecantes, são frequentemente aplicados em área total para o controle das espécies de plantas daninhas predominantes na área (Marochi et al., 2018).

No Brasil, o uso de herbicidas representa 48% do volume total de produtos aplicados na agricultura e o custo médio para o controle de plantas daninhas é cerca de R\$ 62,57 por hectare, dependendo das doses e moléculas utilizadas (Adegas et al., 2017, Sindeveg, 2020). Contudo, a não eficiência no controle de plantas daninhas, bem como, a necessidade de fazer uso de moléculas alternativas de herbicidas devido à resistência de espécies, podem elevar os custos médios para até R\$ 285,98 (Adegas et al., 2017). O desenvolvimento de novas moléculas demanda tempo e custos elevados, estima-se que os gastos mundiais em pesquisa com herbicida estejam entre US \$ 1,2 a US \$ 1,6 bilhão por ano.

Por isso, o desenvolvimento de novas técnicas para identificação de plantas daninhas, contribuem para a adoção de um controle químico mais específico, ou seja, aplicação apenas no local de presença das plantas daninhas. Este tipo de manejo específico pode auxiliar na redução de populações resistentes e viabilizar a aplicação de herbicidas até mesmo em culturas com maior restrição de moléculas disponíveis no mercado (Fennimore e Cutulle, 2019).

Considerar a distribuição espacial e a variabilidade das espécies de plantas daninhas é fundamental para a realização de um manejo racional, aliado aos avanços da agricultura de precisão, que tem possibilitado a adoção de manejos específicos para estas espécies (De Moraes et al., 2008; Reichardt e Jurgens, 2009; Singh et al., 2020). O manejo específico inclui a pulverização dos herbicidas apenas no local onde estão as espécies, além do ajuste de aplicações em função da densidade e composição das plantas daninhas (López-Granados, 2010; Hansen et al., 2013).

A implantação de um manejo específico de plantas daninhas pode ser representada por quatro etapas: 1) monitoramento de plantas daninhas (amostragem ou detecção das espécies); 2) levantamento do histórico e planejamento do manejo; 3) controle de precisão no campo e 4) avaliação dos impactos econômicos e ambientais (López-Granados, 2010).

Contudo, para a adoção de um manejo específico, ainda existem fatores limitantes como o elevado custo de implantação de novas tecnologias, a falta de metodologias eficientes de identificação de plantas daninhas, profissionais na área e informações sobre as suas vantagens (De Moraes et al., 2008). Para López-Granados (2010), a etapa de monitoramento nas lavouras é um dos componentes mais críticos

para a adoção de um manejo específico. Sendo assim, estudos multidisciplinares em conjunto com a implementação de novas tecnologias são fatores que favorecem a adoção de um manejo específico, gerando um retorno rápido e incremental (Reichardt e Jurgens, 2009; Westwood et al., 2017; Klerkx et al., 2019).

### **2.3. Identificação de plantas daninhas**

A adoção de um manejo específico de plantas daninhas é algo que tem ganhado a atenção das empresas e pesquisas no setor agrícola, em função das oportunidades de mercado e do crescente desenvolvimento de novas tecnologias (Van Evert et al., 2017; Klerkx et al., 2019). Além disso, a adoção de manejos específicos para estas espécies a médio e longo prazo, contribuirão com a redução dos custos no manejo e no uso de herbicidas, ao realizar aplicações seletivas e mais eficientes, direcionadas ao alvo (Olsen et al., 2019). Logo, diversas estratégias tem sido desenvolvidas neste âmbito, como, por exemplo, a criação de mapas para aplicação posterior e a detecção de plantas em tempo real, a qual integra sensores aos sistemas de aplicação de herbicidas (Ferreira et al., 2017; Fernández-Quintanilla et al., 2018; Olsen et al., 2019).

Na área de sensoriamento remoto, a distribuição espacial de plantas daninhas dentro de cultivos pode ser detectada e mapeada pela análise de imagens aéreas e de satélite e, também, por sensores proximais, que se referem aos sensores acoplados em barras de pulverização em campo (Wang et al., 2007; Fernández-Quintanilla et al., 2018). O processamento desses mapas consegue distinguir espécies ou grupo de espécies de plantas daninhas, uma vez que, cada espécie possui características próprias, como forma, tamanho, composição química e refletância da folha (Everitt et al., 1992; Bagavathiannan et al., 2019; Fennimore e Cutulle, 2019).

A análise de imagens tem sido bastante explorada para este fim. Ferreira et al., (2017) desenvolveram um software capaz de acertar em 99,5%, plantas daninhas de folhas estreitas e folhas largas, baseados em 15 mil imagens capturadas com o auxílio de um veículo aéreo não tripulável (VANT) na cultura da soja. Olsen et al. (2019) alcançaram um desempenho médio de classificação de até 95,7% na adoção de

modelos de classificação múltipla, baseados em um conjunto de 17.509 imagens de oito espécies de plantas daninhas em pastagens.

No entanto, as técnicas aplicadas ao uso de imagens ainda necessitam de aperfeiçoamento e demandam tempo para a construção de um banco de dados extenso com maior variabilidade e a interpretação dos resultados (López-Granados, 2010; Olsen et al., 2019). Além disso, o uso de imagens não anula a importância da investigação do uso de outras técnicas, como a espectroscopia NIR, visto que, são ferramentas alternativas, capazes de melhorar o desenvolvimento e desempenho dos equipamentos para o manejo específico de plantas daninhas em campo (Olsen et al., 2019; Souza et al., 2020). López-Granados (2010), destaca que todas as técnicas possuem um grau de importância, sendo que, as que adotam o sensoriamento remoto são úteis para identificação e manejo específico baseado em mapas, enquanto as que abordam a detecção proximal é útil para identificação e realização da aplicação em tempo real.

Brown e Noble (2005) revisaram a detecção de plantas daninhas para adoção de manejo específico e concluíram que ainda existe uma grande necessidade de pesquisas em áreas como iluminação artificial, seleção de bandas espectrais, processamento de imagens em diferentes resoluções espaciais e espectrais e imagens capturadas de múltiplas perspectivas.

Os sensores não baseados em imagens, apresentam uma velocidade de processamento melhor na identificação do alvo, quando comparado às análises de imagens (Wang et al., 2007; López-Granadas, 2010; Souza et al., 2020). Já existem aplicações comerciais de alguns sensores sem imagem, para identificação de plantas daninhas em tempo real, como por exemplo, Weedseeker (Trimble Inc., Sunnyvale CA, EUA), AmaSpot (Amazonen-WerkeH. Dreyer GmbH & Co, Hasbergen, Alemanha) e Weed-it (Rometron BV, Steenderen, Holanda) (Fernández-Quintanilla et al., 2018). No entanto, estes sensores disponíveis no mercado só conseguem classificar a ausência ou presença de plantas daninhas, ou seja, possuem uma única função de discriminar as plantas sobre uma determinada superfície.

Identificar diferentes espécies de plantas daninhas em culturas seria uma grande vantagem na hora de recomendar a aplicação local de herbicidas seletivos, visto que estas recomendações são variáveis em função de características biológicas,

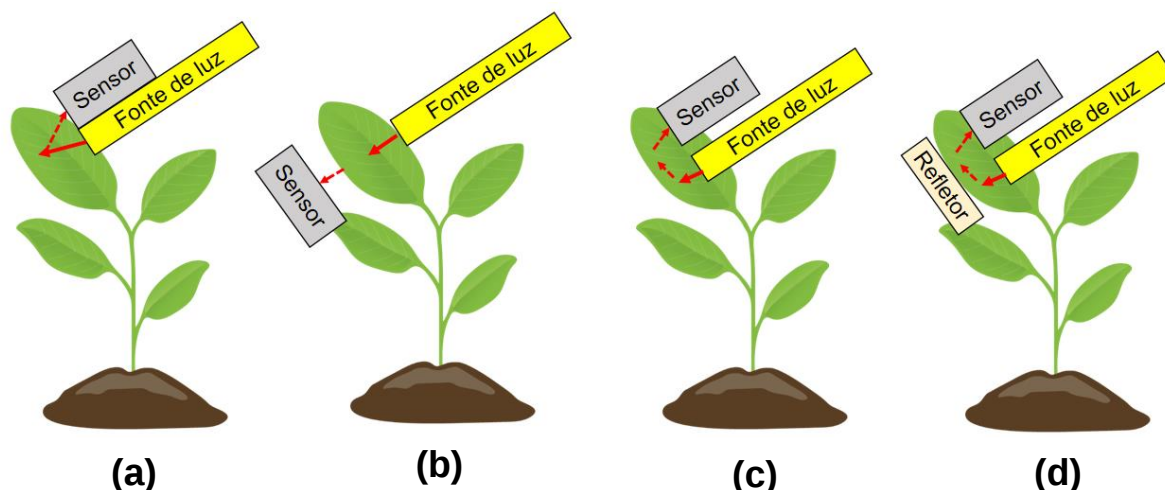
como a espécie e o estágio de desenvolvimento (Orzari et al., 2013; Rodrigues e Almeida, 2018; Ribeiro et al., 2018). Porém, este tipo de manejo automatizado ainda corresponde a um desafio, pois necessita de sistemas de sensores mais sofisticados (Fernández-Quintanilla et al., 2018). Por isso, o desenvolvimento e calibração de modelos para elaboração de sensores específicos é uma das áreas promissoras para o avanço das técnicas usadas para o manejo específico e racional de plantas daninhas (López-Granadas, 2010; Souza et al., 2020). Sendo assim, a espectroscopia de refletância difusa no infravermelho próximo (NIRS), pode ser considerada uma técnica auxiliar, capaz de contribuir com os estudos no manejo específico de plantas daninhas.

#### **2.4. Uso da espectroscopia no infravermelho próximo na agricultura**

Os estudos relacionados ao infravermelho próximo iniciaram-se no século XIX, com o cientista inglês, Frederick William Herschel, que ao estudar as variações de temperatura no espectro do visível (300 a 700 nm) notou que ainda havia variação térmica após a última faixa do visível (Wheeler, 1959; Pasquini, 2003). Herschel chamou estes raios caloríficos de “infrared” (abaixo do vermelho), e a partir dessa descoberta, diversos estudos foram desenvolvidos e as aplicações do infravermelho próximo continuam sendo estudadas para diferentes fins até os dias atuais (Pasquini, 2003; Nicolai, 2007; Cao e Cao, 2020).

A espectroscopia no infravermelho próximo (NIRS) é o resultado da interação entre a radiação e a matéria (amostra) em um intervalo de comprimento de onda aproximado de  $13.300$  a  $4.000\text{ cm}^{-1}$  (750 a 2.500 nm) (Davies, 1998; Pasquini, 2003). Esta técnica consiste na aquisição qualitativa e/ou quantitativa de medidas da radiação que incidem sobre a amostra podendo ser refletida, absorvida ou transmitida (Pasquini, 2003).

A forma de interação da radiação eletromagnética com a amostra pode ser mensurada de distintas formas, e a escolha da melhor forma depende de fatores, como tipo de amostra e o equipamento usado (Pasquini, 2003; Botelho, 2014). As formas mais comuns de mensurar a interação da radiação NIR com a amostra são através da refletância difusa, transmitância, interactância e transfectância (Figura 1).

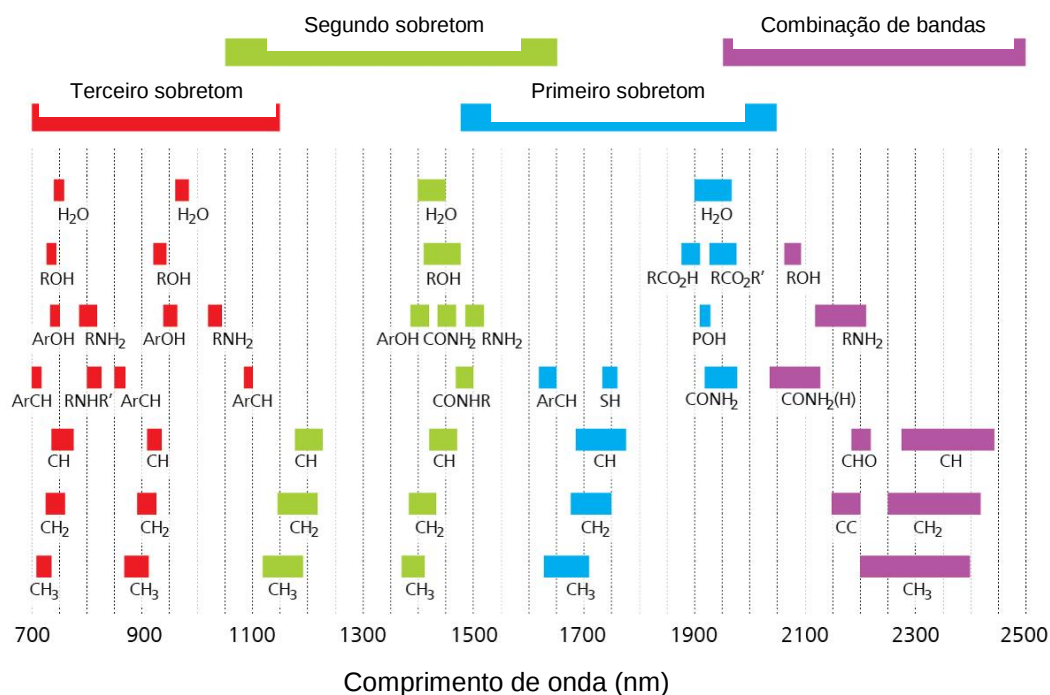


**Figura 1.** Formas de interação da radiação eletromagnética com a superfície amostrada: (a) refletância; (b) transmitância; (c) interactância e (d) transflectância. Fonte: Elaborado pela autora, baseado em Pasquini, (2003).

A combinação da energia da radiação com a diferença de energia entre dois níveis vibracionais provoca uma resposta seletiva do sistema molecular à radiação incidente. Isso significa que em determinado comprimento de onda, uma frequência correspondente é absorvida (Xiaobo et al., 2010). Sendo assim, a intensidade de absorção versus o comprimento de onda, constitui o espectro de absorção de uma substância ou amostra. O principal objetivo do uso da espectroscopia NIR é sondar uma amostra, a fim de adquirir dados qualitativos e/ou quantitativos provenientes da interação de ondas eletromagnéticas no infravermelho próximo com seus componentes. Isso acontece porque a absorção de luz em um dado comprimento de onda é capaz de aumentar a frequência vibracional das ligações entre os átomos, resultando em movimentos harmônicos quantificáveis (Wheeler, 1959; Shenk et al., 2001).

A absorção da luz no infravermelho próximo pode ser aplicada como um método analítico para qualquer amostra que contenha ligações dos grupos químicos de Carbono-Hidrogênio (C-H); Nitrogênio-Hidrogênio (N-H); Enxofre-Hidrogênio (S-H) e Oxigênio-Hidrogênio (O-H) (Pasquini, 2003). As características vibracionais de cada grupo funcional e seus acoplamentos, aparecem em bandas de frequências específicas no espectro (Figura 2) (Xiaobo et al., 2010). A alta variabilidade dessas

características em cada banda faz com que o espectro possa ser considerado uma impressão digital de uma dada amostra (Rohman e Man, 2010; Souza e Poppi, 2012).



**Figura 2.** Grupos funcionais atribuídos aos comprimentos de onda no infravermelho próximo. Fonte: Adaptado, Xiabo et al., (2010).

Entre as vantagens que esta técnica apresenta, as principais estão relacionadas a avaliação de amostras em qualquer estado, a velocidade de aquisição dos resultados e a não necessidade de preparo da amostra. Em resumo, é uma técnica rápida, não destrutiva e de baixo custo (Davies e Grant, 1987; Pasquini, 2003). A principal desvantagem é que NIR não é uma tecnologia autônoma, pois são necessárias calibrações e uso de métodos de referência. Além disso, os espectros são afetados por várias fontes de variabilidade, o que demanda uma grande população de amostras para alcançar modelos robustos de calibração (Workman et al., 2003). Pasquini (2003) recomenda o uso de 50 a 100 amostras, para cada fonte de variável, dependendo da complexidade e variabilidade da matriz que acompanha as amostras.

A espectroscopia NIR envolve abordagens multidisciplinares em química analítica, estatística e engenharia computacional, para conseguir extrair informações químicas significativas de amostras com um alto nível de complexidade e variabilidade

(Workman et al., 2003). Utilizando uma faixa espectral adequada e específica, com NIR, é possível determinar mais de um indicador de qualidade ao mesmo tempo, através da aplicação de métodos quimiométricos de calibração multivariada (Lammertyn et al., 1998; Schmilovitch et al., 2000). Os estudos quimiométricos auxiliam na interpretação dos resultados quantitativos, possibilitando através de um tratamento matemático, estimar novas amostras a partir de modelos criados (Torrent e Barrón, 2008). Além disso, a técnica pode ser aplicada como um método secundário de determinação, que necessita de calibração com base em referências principais para o desenvolvimento e validação de modelos matemáticos (Honorato et al., 2007).

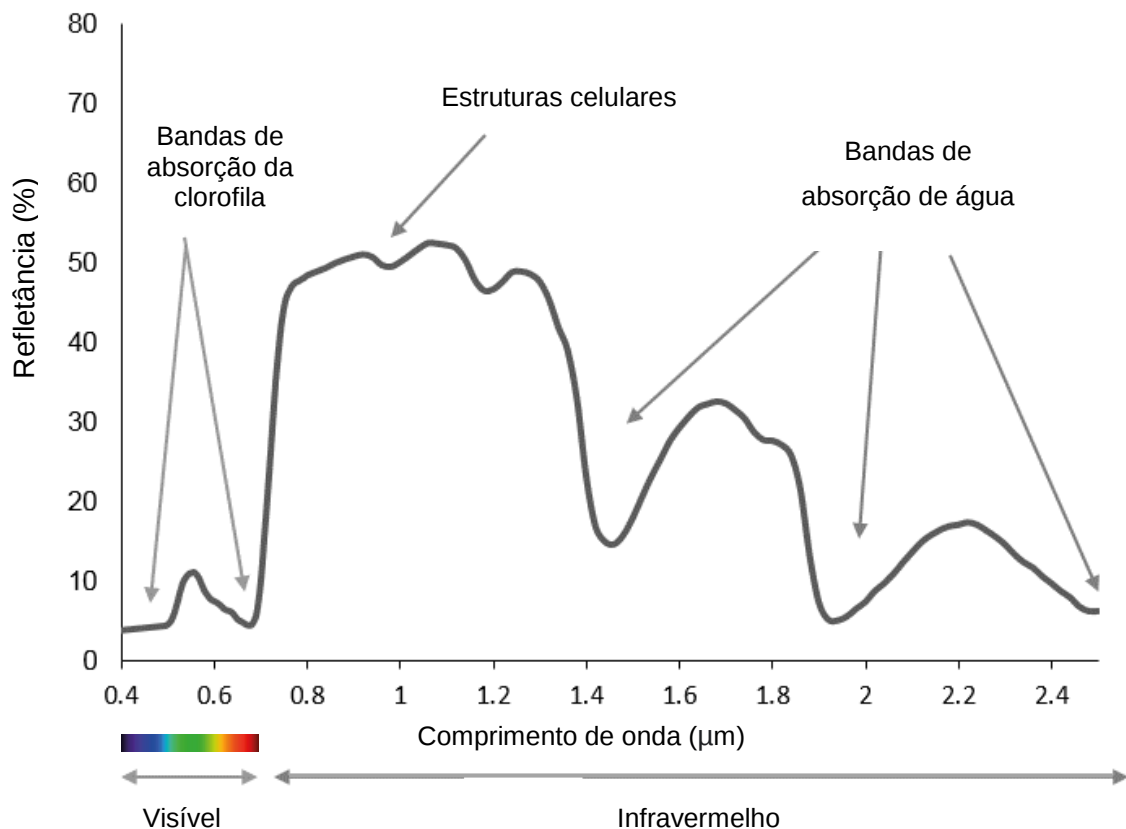
A espectroscopia NIR foi utilizada pela primeira vez na agricultura para avaliar o teor de umidade em grãos (Norris, 1996). Desde então, vem sendo utilizada para análises rápidas de determinação de teores de umidade, nitrogênio, proteína e lipídeos de uma variedade de produtos agrícolas e alimentícios, na análise de solos, na indústria farmacêutica e de petróleo (Davies e Grant, 1987; Pasquini, 2003; Marques et al., 2014; An et al., 2014).

Na agricultura, a técnica tem apresentando bons resultados para determinação dos atributos dos solos (Marques et al., 2014; Demattê et al., 2019), da composição química em forragens (Fontaneli et al., 2004; Berauer et al., 2020), de parâmetros de qualidade em frutos de açaizeiro (Cunha Júnior et al., 2016), em manga (Dos Santos Neto et al., 2017; Sun et al., 2020), em azeite de oliva extravirgem (Sinelli et al., 2010), em leite (Roza-Delgado et al., 2017; Lima et al., 2018); identificação de espécies de cupins (de Azevedo et al., 2019); seleção e discriminação de espécies de madeira (Soares et al., 2017; Snel et al., 2018), cultivares de amendoeira (Borraz-Martinez et al., 2019b), cana-de-açúcar (Steidle Neto et al., 2018) e de plantas daninhas e cana-de-açúcar (Souza et al., 2020).

Sensores multiespectrais elevaram a inspeção visual de lavouras, detectando indicadores precoces de doenças, pragas e necessidades nutricionais, além de monitorar mudanças de carbono orgânico do solo e apresentar novas oportunidades em grãos, incluindo modelagem, melhoramento genético e gerenciamento de risco, estimando com rapidez e precisão características de interesse (Cao e Cao, 2020). Nota-se que a tecnologia tem sido bastante explorada em estudos de predição, caracterização e classificação, e apresenta um alto potencial de aplicação em

modelagem e aprendizado de máquina, contribuindo significativamente com uma agricultura de precisão (An et al., 2014; Demattê et al., 2019; Tian et al., 2020).

Estudos sobre a identificação das características de interesse em plantas cultivadas têm contribuído para a aplicação desta técnica para diferentes fins, principalmente os estudos que retratam as características espectrais de cada espécie (Carvalho, 2005). Em sensoriamento remoto a configuração da curva de refletância espectral de plantas em cada região do espectro (Figura 3), é usada para caracterizar áreas e compreender os atributos biofísicos e bioquímicos da vegetação, bem como, geometria de visualização do detector, condições de iluminação dos arredores e efeitos de fundo (Carvalho, 2005; Prabhakar et al., 2012).



**Figura 3.** Curva de refletância espectral típica de vegetação saudável representando diferentes picos de absorção. Fonte: Adaptado, Prabhakar et al., (2012).

Apesar do potencial promissor do uso de técnicas como a espectroscopia de refletância difusa para classificação e caracterização de espécies, ainda são poucos os estudos que abordam sobre modelos de calibração para espécies de plantas daninhas (Souza et al., 2020). Sendo assim, uma das possibilidades de aplicação da

espectroscopia NIR é no estudo do comportamento espectral de espécies de plantas daninhas, que quando associados às análises multivariadas, poderão auxiliar na calibração de sensores proximais em campo, para realização do manejo específico destas espécies (Olsen et al., 2019; Souza et al., 2020).

## **2.5. Estudos quimiométricos**

Quimiometria é a aplicação de métodos matemáticos, estatísticos e computacionais na investigação, interpretação, classificação e previsão de um conjunto de dados de interesse químico (Pasquini, 2003; Barros Neto et al., 2006). Os principais estudos envolvendo a quimiometria estão relacionados às áreas de planejamento e otimização de experimentos, de reconhecimento de padrões e de calibração multivariada (Barros Neto et al., 2006; Souza e Poppi, 2012).

Na área de planejamento e otimização de experimentos, a quimiometria é aplicada para encontrar as variáveis que mais afetam as características de interesse, assim como a interação entre essas variáveis (Souza e Poppi, 2012). Hanrahan et al. (2008), ressaltam que o uso de técnicas de otimização baseadas em projetos experimentais quimiométricos, foram fundamentais para pesquisadores aprimorarem a capacidade de separação de uma ampla gama de analitos em análises biológicas, ambientais, farmacêuticas, médicas e de tecnologia de alimentos.

Por muitos anos o reconhecimento de padrões químicos era considerado quase um sinônimo da palavra quimiometria e somente a partir de 1980 que estes estudos foram perdendo espaço e tornaram-se uma área menor dentro da quimiometria (Brereton, 2015). O reconhecimento de padrões é usado para analisar um conjunto de dados e identificar agrupamentos ou tendências dentro desses conjuntos (Souza e Poppi, 2012). Nesta área, é usual a aplicação de análises exploratórias e de classificação, supervisionadas ou não, como análise de componente principal (PCA – do inglês: Principal Component Analysis), análise linear discriminante (LDA – do inglês: Linear Discriminant Analysis); análise discriminante por quadrados mínimos parciais (PLS-DA – do inglês: Partial Least Squares Discriminant Analysis) e máquina de vetores suporte (SVM – do inglês: Support Vector Machine) (Brereton, 2015; Wang e Yu, 2015; Steidle Neto et al., 2018).

Por fim, a calibração multivariada, é a responsável por estabelecer modelos que relacionam medidas químicas ou espectrais a uma determinada propriedade da amostra (Souza e Poppi, 2012). Para Ferreira et al. (1999), o termo calibração em quimiometria, refere-se à construção de modelos matemáticos capazes de estabelecer uma relação quantitativa entre os dados de saída (como por exemplo, os espectros NIR) e os parâmetros físicos ou químicos de interesse, previamente determinados por uma técnica analítica de referência. A calibração multivariada é amplamente empregada em estudos de predição, ou seja, usada como uma análise indireta para estimar características de interesse em uma dada amostra (Cunha Júnior et al., 2015; Sans et al., 2018). Para os estudos de calibração, as análises mais usadas são as de regressão, que utilizam às informações da matriz de dados espectrais ( $\mathbf{X}$ ) e da matriz de concentrações ( $\mathbf{Y}$ ), correlacionando-as para obter uma relação direta na fase de calibração (dos Santos Neto, 2018; Sayes et al., 2019).

Atualmente, a espectroscopia NIR é a técnica analítica que mais faz uso da quimiometria e ambas evoluíram de forma conjunta ao longo dos anos, pois cada vez que NIRS amplia sua aplicabilidade e alcança modelos de identificação e quantificação mais robustos, impõe novos desafios à quimiometria, motivando o aprimoramento de suas técnicas (Costa Filho e Poppi, 2002; Pasquini, 2003). Essa interdependência que impulsiona o desenvolvimento da quimiometria associada a NIRS é ainda mais expressiva ao considerarmos os avanços tecnológicos em microcomputadores e em espectrofotômetros, permitindo que milhares de dados sejam obtidos em questão de segundos (Pasquini, 2003; Van Evert et al., 2017; Cao e Cao, 2020).

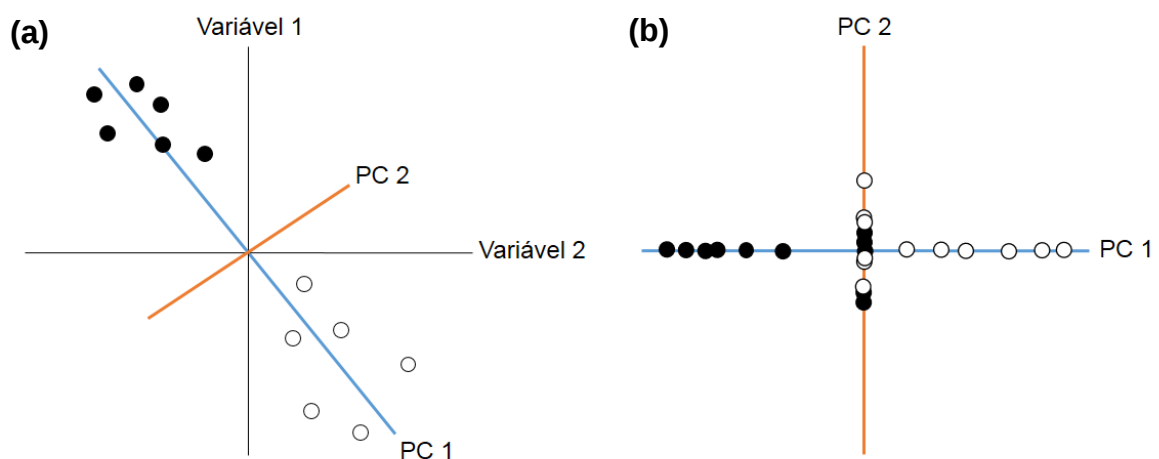
Além dos conhecimentos em estatística, o sucesso da aplicação da quimiometria em NIRS envolve também a etapa de pré-processamento dos dados espectrais (Rinnan et al., 2009; Brereton et al., 2017). Muitos trabalhos aplicam as técnicas de pré-processamento, cujos objetivos principais incluem a eliminação de informações não relevantes e a transformação da matriz de dados, afim de possibilitar uma análise eficiente dos resultados (Souza e Poppi, 2012; Snel et al., 2018; Sayes et al., 2019).

As técnicas de pré-processamento mais utilizadas em NIRS são as que envolvem métodos de correção de dispersão da luz e aplicação de derivadas

espectrais (Rinnan et al., 2009). Entre as técnicas de correção de dispersão destacam-se o uso de correção multiplicativa de sinal (MSC - do inglês: Multiplicative Signal Correction), variação normal padrão (SNV – do inglês: Standard Normal Variate) combinado ou não, com Detrend (DT) e normalização (Rinnan et al., 2009; Soares et al., 2017). Já para o uso de derivadas os pré-processamentos mais comuns são as derivadas de Norris-Williams (NW) e de Savitzky-Golay (SG), que auxiliam na remoção dos efeitos aditivos e multiplicativos dos espectros (Rinnan et al., 2009; Cunha Júnior et al., 2015; Borraz-Martinez et al., 2019).

### 2.5.1 Análise de Componentes Principais (PCA)

A análise de componentes principais (PCA – do inglês: Principal Component Analysis) foi introduzida em 1901, pelo estatístico britânico Karl Person (Person, 1901). A PCA é uma análise capaz de projetar os dados originais multivariados em um espaço de menor dimensão, reduzindo, assim, a dimensionalidade dos dados, sem que as relações entre as amostras sejam afetadas (Ferreira, 2015). Os dados originais são projetados pelos vetores de base, que correspondem aos novos eixos chamados de componentes principais (PC), esses eixos são ortogonais entre si e ordenados conforme o valor da variância explicada, sendo que o primeiro representa a maior variância (Ferreira et al., 1999). A Figura 4 ilustra dois conjuntos de amostras em um espaço bidimensional e os respectivos componentes principais (PC1 e PC2).



**Figura 4.** Representação gráfica de dois grupos de amostras no espaço bidimensional definido pelas Variáveis 1 e 2 e as respectivas componentes principais (a); Representação gráfica das amostras projetadas em PC1 e PC2 (b). Fonte: Adaptado, Ferreira (2015).

A PCA é uma das análises mais importantes na quimiometria e funciona como uma análise base para aplicação de métodos de reconhecimento de padrões, classificação e calibração multivariada (Souza e Poppi, 2012). É considerada como um método de discriminação não supervisionado e é bastante empregada para visualização da estrutura dos dados espectrais, reconhecimento de similaridades entre as amostras e identificação de amostras anômalas, conhecidas como “outliers” (Wu et al., 1997; Souza e Poppi, 2012; Steidle Neto et al., 2019).

### **2.5.2 Análises discriminantes**

As análises discriminantes são métodos de classificação supervisionados usados para o reconhecimento de padrões, quando a propriedade de interesse investigada é categórica ou discreta (Ferreira, 2015). Na análise supervisionada é selecionado um número de amostras para treinamento e outro para validação, utilizando as informações do conjunto de treinamento são desenvolvidos os critérios de reconhecimento que serão usados para a classificação de novas amostras (Pasquini, 2003; Ferreira, 2015; Cunha Júnior et al., 2015). Entre as análises discriminantes supervisionadas, a análise discriminante linear (LDA) e a análise discriminante pelo método dos quadrados mínimos parciais (PLS-DA) são bastante aplicadas em dados espectrais (Ferreira, 2015; Soares et al., 2017; Snel et al., 2018; Cunha Júnior et al., 2015).

A PC-LDA é uma técnica que produz uma série de funções discriminantes lineares ortogonais que maximizam a separação entre os grupos, mas minimizam a variância dentro dos grupos (Naes et al., 2002). Para superar a exigência do LDA de que o número de amostras no conjunto de calibração seja maior do que o número de variáveis, a dimensionalidade dos dados é reduzida usando PCA antes de executar o LDA (Wedding et al., 2019).

A PLS-DA utiliza a regressão PLS para discriminar entre vários grupos, para cada amostra do conjunto de calibração e validação é atribuída uma variável como um valor de referência, indicando a qual classe a amostra pertence (Wedding et al., 2019). Os resultados de um modelo PLS-DA é uma série de valores de previsão para

duas classes, ou seja, valores próximos a 1 para um grupo pré-determinado e próximos a 0 para o restante dos grupos (Ferreira, 2015).

### 2.5.3 Figuras de mérito em análises discriminantes

O bom desempenho de qualquer técnica analítica depende fundamentalmente de dois fatores, o primeiro é a qualidade dos dados amostrados e o segundo é a confiabilidade estatística (Ribeiro et al., 2008). A confiabilidade estatística é a forma de assegurar a aplicabilidade e o alcance de um método, por isso, é necessário definir parâmetros de seleção de análises por meio de figuras de mérito (Ribeiro et al., 2008; Ferreira, 2015). Além disso, é fundamental que existam dois bancos de dados, um para calibração dos modelos e outro para validação, a fim de assegurar a confiabilidade de aplicação do método (Pasquini, 2003; Ribeiro et al., 2008).

As figuras de mérito são, portanto, os indicadores quantitativos do bom desempenho da aplicação das técnicas, sendo que, para análises discriminantes as mais comuns são: taxa de falso positivos (TFP), taxa de falso negativos (TFN), acurácia ou exatidão (EXAT), sensibilidade (SEN) e especificidade ou seletividade (SEL) (Ribeiro et al., 2008; Ferreira, 2015; Soares et al., 2017).

A fim de facilitar a compreensão dos cálculos das figuras de mérito, os resultados são organizados na forma de uma tabela de contingência, também chamada de matriz de classificação ou matriz de confusão (do inglês: *confusion matrix*) (Tabela 1).

**Tabela 1.** Tabela de contingência ou matriz de confusão dos resultados de uma classificação. Fonte: Adaptado, Ferreira (2015).

|                 |   | Classe verdadeira                     |                                       |
|-----------------|---|---------------------------------------|---------------------------------------|
|                 |   | Espécie A                             | Espécie B                             |
| Classe prevista | A | <b>AA</b><br>Verdadeiro Positivo (VP) | <b>BA</b><br>Falso Negativo (FN)      |
|                 | B | <b>AB</b><br>Falso Positivo (FP)      | <b>BB</b><br>Verdadeiro Negativo (VN) |

Na tabela 1, as classes verdadeiras das amostras estão representadas nas colunas, e as classes previstas nas linhas. Quando a classe prevista corresponde a classe verdadeira, significa que o classificador foi efetivo. Quando o classificador não é 100% efetivo, ou seja, não classificou corretamente todas as amostras, significa que houve falsos positivos (FP) e/ou falsos negativos (VN) (Ferreira, 2015).

As figuras de mérito geralmente são expressas em porcentagem e são calculadas conforme as seguintes equações:

$$(1) \text{TFP} = \frac{FP}{FP+VN} 100$$

$$(2) \text{TFN} = \frac{FN}{FN+VP} 100$$

$$(3) \text{EXAT} = \frac{VP+VN}{VP+FN+FP+VN} 100$$

$$(4) \text{SEN} = \frac{VP}{VP+FN} 100$$

$$(5) \text{SEL} = \frac{VN}{FP+VN} 100$$

Além das figuras de mérito apresentadas pelas equações de 1 a 5, existem também outras figuras de mérito, elegíveis conforme o objetivo do classificador, como por exemplo, o índice de Youden ou índice J (Youden, 1959), capaz de considerar as taxas sensibilidade e especificidade ao mesmo tempo.

### 3. REFERÊNCIAS

- Adegas FS, Vargas L, Gazziero DLP, Karam D (2017) **Impacto econômico da resistência de plantas daninhas a herbicidas no Brasil**. Londrina: Embrapa, 11p. (Circular Técnica, 132)
- An X, Li M, Zheng L, Liu Y, Sun H (2014) A portable soil nitrogen detector based on NIRS. **Precision Agriculture** 15(1):3–16.
- Austin DF, Cavalcante PB (1982) Convolvuláceas da Amazônia. **Publicações Avulsas do Museu Paraense Emílio Goeldi**, Vol. 36, 5–134.
- Austin DF, Huáman Z (1996) A Synopsis of Ipomoea (Convolvulaceae) in the Americas. Reviewed work(s): **International Association for Plant Taxonomy (IAPT)**:45.
- Azania AAP, Azania CA, Gravena R, Pavani MCM, Pitelli R (2002) Interferência da palha de cana de açúcar na emergência de espécies de plantas daninhas da família convolvulaceae. **Planta Daninha** 20(2):207–212.
- Bagavathiannan MV, Graham S, Ma Z, Barney JN, Coutts SR, Caicedo AL, De Clerck-Floate R, West NM, Blank L, Metcalf AL, Lacoste M, Moreno CR, Evans JA, Burke I, Beckie H (2019) Considering weed management as a social dilemma bridges individual and collective interests. **Nature Plants** 5(4):343–351.
- Barros Neto B, Scarminio IS, Bruns RE (2006) 25 anos de quimiometria no Brasil. **Química Nova** 29(6):1401–1406.
- Barroso AAM, Ferreira PSH, Martins D (2019) Growth and Development of Ipomoea Weeds. **Planta Daninha** 37:1–9.
- Berauer BJ, Wilfahrt PA, Reu B, Schuchardt MA, Garcia-Franco N, Zistl-Schlingmann M, Dannenmann M, Kiese R, Kühnel A, Jentsch A (2020) Predicting forage quality of species-rich pasture grasslands using vis-NIRS to reveal effects of management intensity and climate change. **Agriculture, Ecosystems and Environment** 296(3):106929.
- 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):1–9.
- Botelho BG (2014) **Desenvolvimento e validação de métodos para análise direta de alimentos usando ferramentas quimiométricas, espectroscopia no infravermelho e imagens digitais**. 136 f. Tese (Doutorado em Química) - Universidade Federal de Minas Gerais, Belo Horizonte.

Braga AF, Barroso AAM, Amaral CL, Nepomuceno MP, Alves PLCA (2018) Population interference of glyphosate resistant and susceptible ryegrass on eucalyptus initial development. **Planta Daninha** 36:e018170148.

Brereton RG (2015) Pattern recognition in chemometrics. **Chemometrics and Intelligent Laboratory Systems** 149:90–96.

Brereton RG, Jansen J, Lopes J, Marini F, Pomerantsev A, Rodionova O, Roger JM, Walczak B, Tauler R (2017) Chemometrics in analytical chemistry—part I: history, experimental design and data analysis tools. **Analytical and Bioanalytical Chemistry** 409(25):5891–5899.

Brown RB, Noble SD (2005) Site-specific weed management: sensing requirements—what do we need to see? **Weed Science** 53(2):252–258.

Campos LHF, Francisco MO, Carvalho SJP, Nicolai M, Christoffoleti PJ (2009) Susceptibility of *Ipomoea quamoclit*, *I. triloba* and *Merremia cissoides* to the herbicides sulfentrazone and amicarbazone. **Planta Daninha** 27(4):831–840.

Cao N, Cao B (2020) NIRS in the contemporary world for food and agriculture. **NIR news** 31(7–8):23–24.

Carvalho AP FDE (2005) Estudo de características foliares de espécies de lenhosas de cerrado e sua relação com os espectros de reflectância. 142 f. Tese (Doutorado) - Instituto de Ciências Biológicas da Universidade de Brasília, Brasília.

Christoffoleti PJ, Borges A, Nicolai M, Carvalho SJP, López-Ovejero RF, Monquero PA (2006) Carfentrazone-ethyl aplicado em pós-emergência para o controle de *Ipomoea* spp. e *Commelina benghalensis* na cultura da cana-de-açúcar. **Planta Daninha** 24(1):83–90.

Correia NM (2016) Controle químico de espécies de corda-de-violão em áreas de cana soca crua colhida nas épocas seca e semiúmida. **Planta Daninha** 34(2):333–343.

Correia NM, Kronka B (2010) Chemical control of plants of the genera *Ipomoea* and *Merremia* in sugarcane. **Planta Daninha** 28(nesp):1143–1152.

Da Costa Filho PA, Poppi RJ (2002) Application of genetic algorithms in the variable selection in mid infrared spectroscopy. Simultaneous determination of glucose, maltose and fructose. **Química Nova** 25(1):46–52.

Cunha Júnior 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.

Cunha Júnior LC, Teixeira GHA, Nardini V, Walsh KB (2016) Quality evaluation of intact açaí and juçara fruit by means of near infrared spectroscopy. **Postharvest Biology and Technology** 112:64–74.

Davies AMC, Grant A (1987) Review: Near infra-red analysis of food. **International Journal of Food Science and Technology** 22:191–207.

Davies T (1998) The history of near infrared spectroscopic analysis: Past, present and future - From sleeping technique to the morning star of spectroscopy. **Analysis** 26(4):17–19.

Demattê JAM, Dotto AC, Paiva AFS, Sato MV, Dalmolin RSD, de Araújo MSB, da Silva EB, Nanni MR, ten Caten A, Noronha NC, Lacerda MPC, de Araújo Filho JC, Rizzo R, Bellinaso H, Francelino MR, Schaefer, C. E. G. R., Vicente, L. E., dos Santos, U. J., de Sá Barretto Sampaio, et al., (2019) The Brazilian Soil Spectral Library (BSSL): A general view, application and challenges. **Geoderma** 354:113793.

Dutra de Moraes V, Agostinetto D, Galon L, Piesanti R, Agostinetto Dirceu, Dutra de Moraes PV, Galon L, Piesanti R (2008) Agricultura de precisão no controle de plantas daninhas precision farming in weed control. **Uruguiana da FZVA** 15(1):1–14.

Everitt JH, Alaniz MA, Escobar DE, Davis MR (1992) Using Remote Sensing to Distinguish Common (Isocoma coronopifolia) and Drummond Goldenweed (Isocoma drummondii). **Weed Science** 40(4):621–628.

Evert FK van, Fountas S, Jakovetic D, Crnojevic V, Travlos I, Kempenaar C (2017) Big Data for weed control and crop protection. **Weed Research** 57(4):218–233.

Felger RS, Austin DF, Devender TRV, Sánchez-Escalante JJ, Costea M (2012) Convolvulaceae of sonora, mexico. I. Convolvulus, cressa, dichondra, evolvulus, ipomoea, jacquemontia, merremia, and operculina. **Journal of the Botanical Research Institute of Texas** 6(11):459–527.

Fennimore SA, Cutulle M (2019) Robotic weeders can improve weed control options for specialty crops. **Pest Management Science** 75(7):1767–1774.

Fernández-quintanilla C, Peña J, Andújar D, Dorado J, Ribeiro A, López-Granados F (2018) Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops? **Weed Research** 58:259–272.

Ferreira, MMC (2015) Quimiometria: conceitos, métodos e aplicações. Campinas: Editora da UNICAMP, 493p.

Ferreira MMC, Antunes AM, Melgo MS, Volpe PLO (1999) Chemometrics i: Multivariate calibration, a tutorial. **Química Nova** 22(5):724-731.

Fontaneli RS, Scheffer-Basso SM, Dürr JW, Appelt JV, Bortolini F, Haubert FA (2004) Predição da composição química de bermudas (Cynodon spp.) pela espectroscopia de reflectância no infravermelho proximal. **Revista Brasileira de Zootecnia** 33(4):838–842.

González A, Amarillo G, Amarillo M, Sarmiento F (2016) Drones aplicados a la agricultura de precisión. **Publicaciones e Investigación** 10:23-37.

Hanrahan G, Montes R, Gomez FA (2008) Chemometric experimental design based optimization techniques in capillary electrophoresis: A critical review of modern applications. **Analytical and Bioanalytical Chemistry** 390(1):169–179.

Hansen KD, Garcia-Ruiz F, Kazmi W, Bisgaard M, La Cour-Harbo A, Rasmussen J, Andersen H J (2013) An autonomous robotic system for mapping weeds in fields. **IFAC Proceedings Volumes** 8:217–224. (IFAC-PapersOnline)

Honorato FA, Neto BDB, Martins MN, Galvão RKH, Pimentel MF (2007) Calibration transfer in multivariate methods. **Quimica Nova** 30(5):1301–1312.

Kissmann KG, Groth D (1999) Convolvulaceae Juss. In: **Plantas infestantes e nocivas**. 3.ed. São Paulo: BASF, p.673-693.

KISSMAN, KG, GROTH D. *Convolvulaceae Juss.* In: **Plantas infestantes e nocivas** 3.ed. São Paulo: BASF, p. 617-754.

Klerkx L, Jakku E, Labarthe P (2019) 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** 90–91(11):100315.

Kowalski BR (1975) Chemometrics: Views and Propositions. **Journal of Chemical Information and Computer Sciences** 15(4):201–203.

Kuva MA, Pitelli RA, Salgado TP, Alves PLCA (2007) Fitossociologia de comunidades de plantas daninhas em agroecossistema cana-crua. **Planta Daninha** 25(3):501–511.

Lammertyn J, Ooms K, Smedt V, Baerdemaeker J (1998) Non-destructive measurement of acidity, soluble solids, and firmness of jonagold apples using nir-spectroscopy. **American Society of Agricultural Engineers** 41(4):1089–1094.

Lima GF, Andrade SAC, da Silva VH, Honorato FA (2018) Multivariate Classification of UHT Milk as to the Presence of Lactose Using Benchtop and Portable NIR Spectrometers. **Food Analytical Methods** 11(10):2699–2706.

López-Granados F (2011) Weed detection for site-specific weed management: Mapping and real-time approaches. **Weed Research** 51(1):1–11.

Lorenzi H (1991) **Plantas daninhas do Brasil: terrestres, aquáticas, parasitas, tóxicas e medicinais**. Nova Odessa: Instituto Plantarum. v.2,

Lorenzi H, Matos, FJA (2002) **Plantas medicinais no Brasil/ Nativas e exóticas**. Nova Odessa: Instituto Plantarum, 512 p.

Marochi A, Ferreira A, Takano HK, Oliveira Junior RS, Ovejero RFL (2018) Managing glyphosate-resistant weeds with cover crop associated with herbicide rotation and mixture. **Ciencia e Agrotecnologia** 42(4):381–394.

Martinelli LA, Naylor R, Vitousek PM, Moutinho P (2010) Agriculture in Brazil: impacts, costs, and opportunities for a sustainable future. **Current Opinion in Environmental Sustainability** 2(5-6):431-438.

Marques Jr. J, Siqueira DS, Camargo LA, Teixeira DDB, Barrón V, Torrent J (2014) Magnetic susceptibility and diffuse reflectance spectroscopy to characterize the spatial variability of soil properties in a brazilian haplustalf. **Geoderma** 219–220, 63–71.

Monquero PA, Silva PV, Hirata ACS, Martins FRA (2011) Monitoramento do banco de sementes de plantas daninhas em áreas com cana-de-açúcar colhida mecanicamente. **Planta Daninha** 29(1):107–119.

Naes T, Isaksson T, Fearn T, Davies T, A (2002) **A user-friendly guide to multivariate calibration and classification**. Chichester: NIR. v.6.

Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J (2007) Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. **Postharvest Biology and Technology** 46(2):99–118.

Norris KH (1996) History of Near Infrared Spectroscopy. **Journal of Near Infrared Spectroscopy** 37(4):31–37.

Oerke EC, Dehne HW (2004) Safeguarding production - Losses in major crops and the role of crop protection. **Crop Protection** 23(4):275–285.

Olsen A, Konovalov DA, Philippa B, Ridd P, Wood JC, Johns J, Banks W, Girgenti B, Kenny O, Whinney J, Calvert B, Azghadi MR, White RD (2019) DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning. **Scientific Reports** 9(1):1–12.

Orzari I, Monquero PA, Reis FC, Sabbag RS, Hirata ACS (2013) Germinação de espécies da família convolvulaceae sob diferentes condições de luz, temperatura e profundidade de semeadura. **Planta Daninha** 31(1):53–61.

Owen MDK, Beckie HJ, Leeson JY, Norsworthy JK, Steckel LE (2015) Integrated pest management and weed management in the United States and Canada. **Pest Management Science** 71(3):357–376.

Pagnoncelli FB, Trezzi MM, Brum B, Vidal RA, Portes ÁF, Scalcon EL, Machado A (2017) Morning glory species interference on the development and yield of soybeans. **Bragantia** 76(4):470–479.

Pannacci E, Lattanzi B, Tei F (2017) Non-chemical weed management strategies in minor crops: A review. **Crop Protection** 96:44–58.

Pasquini C (2003) Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society** 14(2):198–219.

Pearson K (1901) On lines and planes of closest fit to systems of points in space. **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science** 2(11):559-572.

Pitelli RA (1985) Interferência de plantas daninhas em culturas agrícolas. **Informe Agropecuário** 11:16-27.

Pitelli RA (1987) Competição e controle das plantas daninhas em áreas agrícolas. **Série técnica IPEF** 4(12):1-24.

Piza CST, Nepomuceno MP, Alves PLDCA (2016) Period prior to interference of morning glory in sugarcane. **Científica** 44(4):543.

Prabhakar M, Prasad YG, Rao MN (2012) Remote sensing of biotic stress in crop plants and its applications for pest management. In: Venkateswarlu B, Shanker AK, Shanker C, Maheswari M. **Crop stress and its management: perspectives and strategies**. Berlin: Springer, p.517-545.

Reichardt M, Jürgens C (2009) Adoption and future perspective of precision farming in Germany: Results of several surveys among different agricultural target groups. **Precision Agriculture** 10(1):73–94.

Ribeiro FADL, Ferreira MMC, Morano SC, Silva LRD, Schneider RP (2008) Planilha de validação: uma nova ferramenta para estimar figuras de mérito na validação de métodos analíticos univariados. **Química Nova** 31(1):164-171.

Ribeiro NM, Torres BA, Ramos SK, dos 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(7):1090–1098.

Rinnan Å, Berg F van den, Engelsen SB (2009) Review of the most common pre-processing techniques for near-infrared spectra. **TrAC - Trends in Analytical Chemistry** 28(10):1201–1222.

Rohman A, Man YBC (2010) Fourier transform infrared (FTIR) spectroscopy for analysis of extra virgin olive oil adulterated with palm oil. **Food Research International** 43(3):886–892.

Roza-Delgado B, Garrido-Varo A, Soldado A, González Arrojo A, Cuevas Valdés, M, Maroto F, Pérez-Marín D (2017) Matching portable NIRS instruments for in situ monitoring indicators of milk composition. **Food Control** 76:74–81.

Saeys W, Nguyen Do Trong N, Van Beers R, Nicolai BM (2019) Multivariate calibration of spectroscopic sensors for postharvest quality evaluation: A review. **Postharvest Biology and Technology** 158(Jul.):110981.

Sans S, Ferré J, Boqué R, Sabaté J, Casals J, Simó J (2018) Determination of chemical properties in ‘calçot’ (*Allium cepa* L.) by near infrared spectroscopy and multivariate calibration. **Food Chemistry** 262(2):178–183.

Santos CG, Martins GLM, Silva PA, Zoz T (2017) Custo De Produção De Soja Transgênica Em Sistema De Semeadura Direta Em São Desidério-Ba. **Journal of Neotropical Agriculture** 4(2):96–101.

Santos Ferreira A dos, Matte Freitas D, Gonçalves da Silva G, Pistori H, Theophilo Folhes M (2017) Weed detection in soybean crops using ConvNets. **Computers and Electronics in Agriculture** 143(11):314–324.

dos Santos Neto JP (2018) **Otimização Das Condições De Armazenamento Sob Atmosfera Controlada De Mangas “Palmer” Destinadas À Exportação**. 150 f. Tese (Doutorado em Agronomia) – Unesp, Jaboticabal.

dos Santos Neto JP, de Assis MWD, Casagrande IP, Cunha Júnior LC, de Almeida Teixeira GH (2017) Determination of ‘Palmer’ mango maturity indices using portable near infrared (VIS-NIR) spectrometer. **Postharvest Biology and Technology** 130(5):75–80.

Schmilovitch Z, Mizrach A, Hoffman A, Egozi H, Fuchs Y (2000) Determination of mango physiological indices by near-infrared spectrometry. **Postharvest Biology and Technology** 19(3):245–252.

Shenk JS, Workman JJ, Westerhaus MO (2001) Application of NIR spectroscopy to agricultural products. **Practical Spectroscopy Series** 27:419-474.

Silva IAB, Kuva MA, Alves PLC, Salgado TP (2009) Interferência de uma comunidade de plantas daninhas de Ipomoea heredifolia na cana-soca. **Planta Daninha** 27(2):265–272.

Sinelli N, Cerretani L, Egidio VD, Bendini A, Casiraghi E (2010) Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity. **Food Research International** 43(1):369–375.

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.

Snel FA, Braga JWB, da Silva D, Wiedenhoeft AC, Costa A, Soares R, Coradin VTR, Pastore TCM (2018) Potential field-deployable NIRS identification of seven Dalbergia species listed by CITES. **Wood Science and Technology** 52(5):1411–1427.

Soares LF, Da Silva DC, Bergo MCJ, Coradin VTR, Braga JWB, Pastore TCM (2017) Avaliação de espectrômetro nir portátil e pls-da para a discriminação de seis espécies similares de madeiras Amazônicas. **Química Nova** 40(4):418–426.

Souza AM, Poppi RJ (2012) Teaching experiment of chemometrics for exploratory analysis of edible vegetable oils by mid infrared spectroscopy and principal component analysis: A tutorial, part I. **Química Nova** 35(1):223–229.

Souza MF, Amaral LR, Oliveira SRM, Coutinho MAN, Ferreira Netto C (2020) Spectral differentiation of sugarcane from weeds. **Biosystems Engineering** 190:41–46.

Stefanović S, Austin DF, Olmstead RG (2003) Classification of Convolvulaceae: A Phylogenetic Approach. **Systematic Botany** 28(4):791–806.

Steidle Neto AJ, Lopes DC, Toledo JV, Zolnier S, Silva TGF (2018) Classification of sugarcane varieties using visible/near infrared spectral reflectance of stalks and multivariate methods. **Journal of Agricultural Science** 156(4):537–546.

Sun X, Subedi P, Walsh KB (2020) Achieving robustness to temperature change of a NIRS-PLSR model for intact mango fruit dry matter content. **Postharvest Biology and Technology** 162(1):111117.

Tian XY, Aheto JH, Bai JW, Dai C, Ren Y, Chang X (2020) Quantitative analysis and visualization of moisture and anthocyanins content in purple sweet potato by Vis–NIR hyperspectral imaging. **Journal of Food Processing and Preservation** (6):1–16.

Torrent J, Barrón V (2008) Diffuse reflectance spectroscopy. **Methods of Soil Analysis - Mineralogical Methods** 5:367–385.

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(3):276–285.

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(5):277–284.

Wedding BB, Wright C, Grauf S, Gadek P, White RD (2019) The application of FT-NIRS for the detection of bruises and the prediction of rot susceptibility of ‘Hass’ avocado fruit. **Journal of the Science of Food and Agriculture** 99(4):1880–1887.

Westwood JH, Charudattan R, Duke SO, Fennimore SA, Marrone P, Slaughter DC, Swanton C, Zollinger R (2018) Weed management in 2050: perspectives on the future of weed science. **Weed Science** 66(3):275–285.

Wheeler OH (1959) Near Infrared Spectra of Organic Compounds. **Chemical Reviews** 59(4):629–666.

Workman J, Koch M, Veltkamp DJ (2003) Process analytical chemistry. **Analytical Chemistry** 75(12):2859–2876.

Wu W, Massart DL, De Jong S (1997) The kernel PCA algorithms for wide data. Part I: Theory and algorithms. **Chemometrics and Intelligent Laboratory Systems** 36(2):165–172.

Xiaobo Z, Jiewen Z, Povey MJW, Holmes M, Hanpin M (2010) Variables selection methods in near-infrared spectroscopy. **Analytica Chimica Acta** 667(1–2):14–32.

Youden WJ (1950) Index for rating diagnostic tests. **Cancer** 3(1):32–35.

## CAPÍTULO 2 - Discriminação de espécies de corda-de-viola através do uso de espectroscopia NIR e métodos multivariados

**RESUMO** - A ocorrência de espécies de plantas daninhas é um dos principais fatores que limitam a produtividade de culturas agrícolas. O manejo destas espécies é realizado comumente com herbicidas, recomendados para aplicação em área total e em dose única, sem considerar a distribuição espacial. Os estudos de novas técnicas para identificação destas espécies contribuem para o desenvolvimento de sensores proximais, que, no futuro, poderão ser acoplados às máquinas, otimizando a realização de um manejo específico. Sendo assim, o objetivo deste trabalho foi utilizar a espectroscopia no infravermelho próximo (NIRS) para discriminar três espécies de corda-de-viola através de técnicas de análise multivariada. Os espectros foram coletados nas folhas das três espécies em estágio vegetativo (até cinco folhas), através de um espectrômetro NIR de bancada na faixa espectral de 4.000 a 10.000  $\text{cm}^{-1}$ . Os modelos de discriminação foram selecionados através do estudo da acurácia, sensibilidade, especificidade e índice de Youden, observados no conjunto de dados de validação ( $n=135$ ). A técnica de NIRS associada às análises multivariadas, linear discriminante (PC-LDA) e quadrado mínimos parciais (PLS-DA), pode ser utilizada para discriminação de espécies de corda-de-viola. Os melhores resultados foram observados quando se realizou o estudo de faixas espectrais associadas ao uso de pré-processamentos. Os modelos mais assertivos apresentaram acurácias de validação de 99,26, 98,52 e 98,73%, para *Ipomoea hederifolia*, *I.nil* e *Merremia aegyptia*, respectivamente.

**Palavras-chave:** assinatura espectral, *Ipomoea hederifolia*, *Ipomoea nil*, manejo específico de plantas daninhas, *Merremia aegyptia*

**ABSTRACT** – The occurrence of weeds is one of the main factors limiting agricultural productivity. Studies on new techniques for the identification of these species contribute to the development of proximal sensors, which, in the future, might be coupled to machines, optimizing the performance of specific management practices. Thus, the objective of this work was to use near-infrared spectroscopy (NIR) and multivariate analysis to discriminate three morning glory species. The NIR spectra were collected from the leaves of the three species when they presented up to five leaves (vegetative stage) within the spectral band of 4,000 to 10,000  $\text{cm}^{-1}$ . The discrimination models were selected according to accuracy, sensitivity, specificity, and the Youden's index, analyzed within a group validation data set ( $n = 135$ ). The best results occurred when the selection of spectral bands associated with the use of pre-processing was performed. It was possible to obtain an accuracy of 99.26, 98.52, and 98.73% for *Ipomoea hederifolia*, *Ipomoea nil*, and *Merremia aegyptia*, respectively. NIR spectroscopy associated with principal component analysis and linear discriminant analysis (PC-LDA), or associated with partial least squares regression with discriminant analysis (PLS-DA), can be used to discriminate morning glory species.

**Keywords:** *Ipomoea hederifolia*, *Ipomoea nil*, *Merremia aegyptia*, specific weed management, spectral signature

## 1. INTRODUÇÃO

A presença de plantas daninhas em áreas agrícolas é um dos principais fatores que limitam a produtividade. Atualmente, a aplicação de herbicidas é o método mais utilizado para o controle destas espécies, por apresentar uma boa relação entre custo e benefício. No entanto, a aplicação destes, quando mal realizada, propicia o aumento da resistência das plantas a herbicidas e quando realizada em área total, gera desperdícios, aumentando custos e risco ambiental. (López-Granados, 2011; Braga et al., 2018).

Embora existam alguns sensores para a identificação de plantas daninhas disponíveis no mercado, a aplicabilidade destes ainda é limitada, não podendo distinguir as espécies infestantes da cultura de interesse (Dammer and Wartenberg, 2007; López-Granados, 2011; de Souza et al., 2020). Sendo assim, o aprimoramento de técnicas que auxiliem a adoção do manejo específico de plantas daninhas é importante para alcançar uma agricultura cada vez mais sustentável, visto que o ideal é que estas aplicações sejam variáveis em função das espécies e do estágio de desenvolvimento (López-Granados, 2011; de Souza et al., 2020; Rodrigues e Almeida, 2018).

Para isso, a espectroscopia no infravermelho próximo (NIRS), pode ser uma alternativa na identificação de espécies de plantas daninhas a campo, uma vez que esta técnica tem sido bastante utilizada no monitoramento e avaliação da qualidade de diversos produtos agrícolas (Nicolai et al., 2007; Jamshidi e Mohajerani, 2016). Sua capacidade de diferenciar conjuntos específicos de moléculas, quando associada às análises multivariadas, permite separar as amostras em lotes específicos através da “impressão digital” que cada amostra possui e também associar isso às informações físicas e anatômicas das espécies (Soares et al., 2017).

Esta técnica já foi utilizada com êxito para separar lotes de espécies em viveiros, para discriminação de cultivares não diferenciáveis visualmente e também para a identificação de espécies de madeiras nativas (Wang e Yu, 2015; Soares et al., 2017; Steidle Neto et al., 2018; Borraz-Martínez et al., 2019). Além disso, a alta velocidade de aquisição dos espectros NIR em combinação com técnicas de calibração multivariada, tem possibilitado a classificação on-line (em tempo real) de

frutas e vegetais com base em atributos de qualidade, que somente pela aparência também não seria possível detectar diferenças (Cunha Júnior et al., 2016; de Souza et al., 2020).

Apesar do potencial de uso da espectroscopia NIR para a classificação e caracterização de espécies, ainda não existem modelos de calibração e bancos de dados espectrais desenvolvidos para espécies de plantas daninhas. Sendo assim, a aplicação desta técnica em plantas daninhas pode ser considerada uma ferramenta em potencial na diferenciação no campo e discriminação destas espécies. Para isso, estudos quimiométricos são fundamentais, visto que as informações espectrais são complexas e necessitam de modelos matemáticos para relacionar os espectros às características de interesse (Pasquini, 2003). Estes estudos, aliados a um elevado número de amostras, permitirão entender algumas propriedades químicas que posteriormente irão auxiliar na diferenciação das espécies de maneira mais rápida e em níveis cada vez mais específicos (López-Granados, 2011; Carvalho et al., 2013; de Souza et al., 2020).

Desse modo, o objetivo deste trabalho foi discriminar três espécies de plantas daninhas, pertencentes a família Convolvulaceae, conhecidas popularmente como corda-de-viola, usando NIRS e análises multivariadas.

## **2. MATERIAL E MÉTODOS**

### **2.1. Material vegetal**

Sementes de *Ipomoea hederifolia*, *Ipomoea nil* e *Merremia aegyptia* foram semeadas em vasos plásticos com capacidade para 0,5 litros preenchidos com areia de rio grossa lavada e peneirada, deixando-se uma planta por vaso. Todas as plantas foram cultivadas a céu aberto, sob coordenadas 21°14'38.6" S e 48°17'56.3" O. Foram cultivados 150 vasos por espécie, totalizando 450 vasos, e cada vaso correspondeu a uma parcela experimental (n=450). Os vasos foram irrigados diariamente com solução nutritiva completa de Hoagland e Arnon (1938). A fim de promover a adaptação das plântulas à solução nutritiva, foi realizado o aumento gradual e semanal da força iônica da solução. Na primeira semana, a solução nutritiva foi

ajustada para uma concentração de 25%, na segunda semana por 50%, consecutivamente até que atingisse 100% de força iônica. O período de condução das espécies foi de março a abril de 2019, e o clima da região segundo a classificação de Köppen é Aw (Rolim et al., 2007).

Após a coleta dos espectros NIR, foi realizada a caracterização das espécies pelo número de folhas e biomassa fresca e seca da parte aérea, a fim de mensurar o teor de água nas amostras. Para determinação da biomassa seca, as plantas foram acondicionadas em sacos de papel e secas em estufa com circulação forçada de ar, com temperatura de até 40 grau Celsius até massa constante.

## **2.2. Obtenção dos espectros de refletância**

Os espectros NIR foram obtidos através do acessório NIR Fiber Optic Probe acoplado ao espectrofotômetro FT-NIR Spectrum 100N (Perkin Elmer. Shelton. Estados Unidos da América). Os espectros de refletância difusa foram adquiridos na forma de  $\log(1/R)$ , onde R corresponde à refletância, na faixa espectral de 4.000-10.000  $\text{cm}^{-1}$  (1.000 -2.500 nm), com 32 scans, a uma resolução espectral de 4  $\text{cm}^{-1}$  e com intervalo de 2  $\text{cm}^{-1}$ . No momento da coleta dos espectros NIR, a temperatura do ambiente foi mantida a  $23^{\circ}\text{C} \pm 1$ , permanecendo constante durante todo o processo de aquisição espectral.

As curvas espectrais foram obtidas do terço médio da parte adaxial da última folha totalmente expandida, evitando-se as áreas com nervuras. Cada planta foi amostrada posicionando a folha na saída do feixe de luz sob um fundo preto (de EVA)), utilizado para evitar a entrada de luz externa.

## **2.3. Desenvolvimento de modelos de classificação**

Os dados espectrais foram analisados usando o programa Unscrambler ® versão 10.3 (CAMO. Oslo. Noruega). Similarmente, este programa foi utilizado para fazer os pré-processamentos dos espectros NIR, ou seja. SNV (*Standard Normal Variate*), Detrend, SNV+Detrend, 1ª derivada de Savitzky–Golay (1SG) agrupado em 6 e 6 pontos, e SNV + 1SG. A análise de componente principal (PCA) foi realizada com todo o espectro NIR, em faixas selecionadas, sem e com os pré-processamentos, para medir o nível da variação espectral entre as amostras.

Para discriminar as espécies foram desenvolvidos modelos de classificação usando a análise discriminante linear de componentes principais (PC-LDA) e análise discriminante por mínimos quadrados parciais (PLS-DA). Para isso, as três espécies foram consideradas individualmente (*I. hederifolia*, *I. nil* e *M. aegyptia*) e também em diferentes faixas espectrais, definidas conforme os valores dos loadings da PCA com o espectro bruto, sem pré-processamento. Os comprimentos de onda selecionados para o estudo foram: 1) 10.000-4.000  $\text{cm}^{-1}$  (1.000-2.500 nm); 2) 7.460-4.240  $\text{cm}^{-1}$  (1.340-2.358 nm); 3) 7.460-6.800  $\text{cm}^{-1}$  (1.340-1.471 nm); 4) 5.540-4.240  $\text{cm}^{-1}$  (1.805-2.358 nm) e 5) 7.460-6.800+5.540-4.240  $\text{cm}^{-1}$  (1.340-1.471+1.805-2.358 nm).

Foram utilizados três métodos de separação dentro da PC-LDA: distância linear, quadrática e distância de Mahalanobis. O método de Mahalanobis foi mantido, por apresentar o menor erro de classificação. Foram utilizados 10 componentes principais (PCs) e o método de validação cruzada aleatória, com 31 segmentos. Na análise de PLS-DA, para cada amostra do conjunto de calibração e validação, foi atribuída uma variável como um valor de referência indicando a qual classe a amostra pertence. Foi atribuído o valor numérico de 1 a uma classe e o valor 0 a outra classe, foi definido também um valor limite, ou ponto de corte, de 0.5 para determinar a atribuição de classe (Soares et al., 2017).

Para dividir os dados espectrais original ( $n=450$ ) em conjunto de calibração ( $n=315$ ) e de validação ( $n=135$ ), empregou-se o algoritmo de Kennard-Stone (Kennard e Stone, 1969), que baseia-se em uma abordagem sequencial determinística e busca selecionar amostras uniformemente distribuídas no espaço de previsão. Após a validação interna (calibração), a capacidade de previsão dos modelos foi testada usando as amostras que não foram incluídas neste conjunto, ou seja, realizou-se a validação externa com 30% do conjunto total dos dados, selecionando os melhores pré-processamentos para cada espécie em função dos valores de acurácia na validação e especificidade, além das faixas espectrais.

O desempenho dos modelos foi avaliado pelos valores de acurácia para o banco de dados de calibração e de validação e, também, pelos índices de sensibilidade, especificidade e pelo índice de Youden (1950), calculados conforme as seguintes equações:

(1) Acurácia (%) =  $100 - \frac{y}{N} \times 100$  , onde y = total de amostras classificadas incorretamente em um total de N amostras;

(2) Sensibilidade (%) =  $\left( \frac{TP}{TP+FN} \right) \times 100$ , onde TP= verdadeiro positivo e FN= falso negativo;

(3) Especificidade (%)=  $\left( \frac{TN}{TN+FP} \right) \times 100$ , onde TN= verdadeiro negativo e FP= falso positivo;

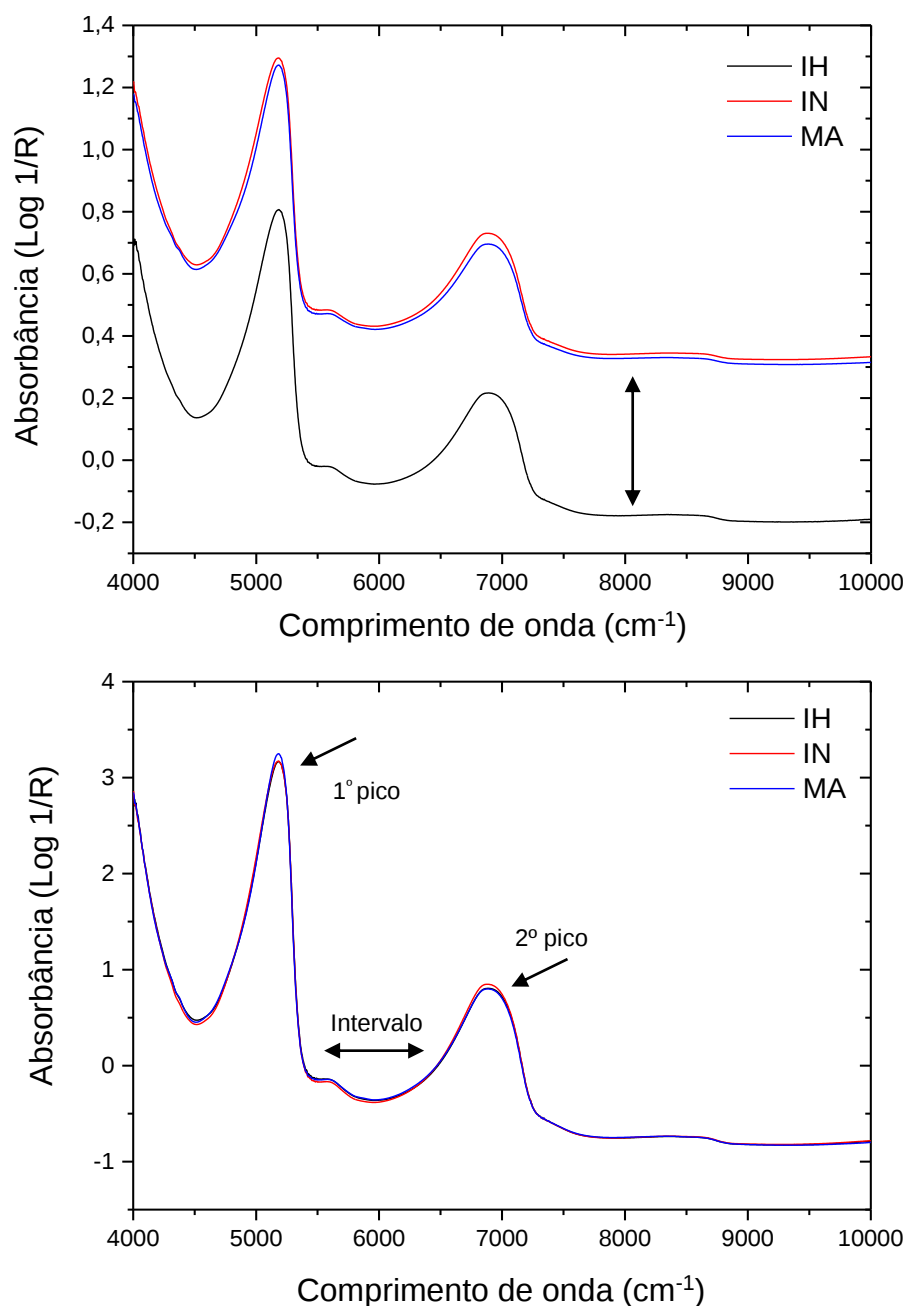
(4) Índice de Youden- J (%) = SENS – (100 – SPEC), onde SENS= sensibilidade e SPEC = especificidade.

### 3. RESULTADOS E DISCUSSÃO

#### 3.1. Caracterização e espectro NIR

Não houve diferença significativa ( $p < 0.05$ ) para os dados de caracterização e teores de água nas três espécies de corda-de-viola no momento da coleta dos espectros. O teor de água médio nas folhas das espécies foi de  $81\% \pm 2$ . Estes resultados demonstram que a umidade das folhas frescas não foi a principal variação responsável por discriminar as espécies através dos espectros NIR (Nicolaï et al., 2007; Soares et al., 2017).

A média dos espectros NIR das folhas das três espécies de corda-de-viola estão apresentados na Figura 1. Nos espectros sem pré-processamento (Figura 1a) pode-se observar um deslocamento do espectro médio da espécie *I. hederifolia* em relação às demais. Nos espectros brutos, sem pré-processamento, geralmente ocorre uma variação da linha de base, causada principalmente pela presença de espalhamento aditivo e multiplicativo da luz. Essa variação é principalmente resultado de efeitos físicos devido a diferenças na superfície das amostras ou a variações no momento da coleta dos espectros (Pasquini, 2003; Snel et al., 2018).



**Figura 2.** Dados de absorbância (Log 1/ Refletância) do espectro todo, sem pré-processamento (a) e com pré-processamento SNV (b) de *I.hederifolia* (IH), *I.nil* (IN) e *M.aegyptia* (MA).

Embora não tenha apresentado diferença estatística, as espécies apresentavam de 2 a 5 folhas. Portanto, entende-se que as folhas jovens podem ter ocasionado um maior escape de luz em relação às mais velhas, alterando assim os valores de absorbância (Figura 1a). Desta forma, utilizou-se o pré-processamento dos

espectros NIR para a correção dos problemas relacionados ao escape de luz, deslocamento da linha de base e a dispersão multiplicativa. (Cunha Junior et al., 2015; Snel et al., 2018). Como pode-se observar na Figura 1a, a diferença entre os espectros NIR se deu principalmente a um acentuado desvio da linha de base entre as espécies e após aplicar um dos pré-processamentos, SNV (Figura 1b), essa diferença foi corrigida. O mesmo comportamento de ajuste dos espectros foi observado para os outros pré-processamentos selecionados.

Os picos de um espectro são as regiões que possuem as maiores informações e foram observados dois principais picos nas faixas de  $4.240\text{--}5.540\text{ cm}^{-1}$  e  $6.800\text{--}7.460\text{ cm}^{-1}$  (Figura 1b). A faixa de  $7.500\text{--}10.000\text{ cm}^{-1}$  é a região com menores informações sobre as amostras. Este mesmo comportamento espectral também foi observado em *Prunus dulcis* (Borraz-Martínez et al., 2019).

Através do estudo das bandas espectrais de NIR é possível identificar os principais grupos químicos referentes a determinadas faixas do espectro (Pasquini, 2003; Cunha Junior et al., 2015). Os picos observados na região de  $5.000\text{ e }7.000\text{ cm}^{-1}$  corresponderam às vibrações O-H; neste caso, a água é a principal responsável por essas bandas. Porém, essas regiões também podem representar grupos de compostos com radicais  $\text{CO}_2\text{H}$  e  $\text{CO}_2\text{R}'$  (Xiaobo et al., 2010).

A principal dificuldade dos estudos com espectroscopia NIR em produtos vegetais frescos é que, geralmente, as regiões com mais informações para determinados constituintes também são as regiões que sobrepõem aos picos relativos à presença de água, dificultando uma análise quantitativa e a seleção de um único comprimento de onda para um determinado composto (Cunha Junior et al., 2015; Jintao et al., 2018; Borraz-Martínez et al., 2019).

As espécies da família Convolvulaceae apresentam uma ampla gama de alcaloides e glicosídeos, que podem variar em ocorrência de gênero, espécie e estágio de desenvolvimento (Jenett-Siems et al., 1998; Meira et al., 2012; Carvalho et al., 2013). Alguns alcaloides foram identificados em *Coptidis rhizoma* por meio da espectroscopia NIR usando a regressão por quadrados mínimos parciais (PLSR) nos comprimentos de onda de  $4.249\text{--}7.502\text{ cm}^{-1}$ ;  $4.596\text{--}7.502\text{ cm}^{-1}$  e  $4.249\text{--}6.100\text{ cm}^{-1}$  (Jintao et al., 2018). Estas regiões correspondem às mesmas regiões dos picos

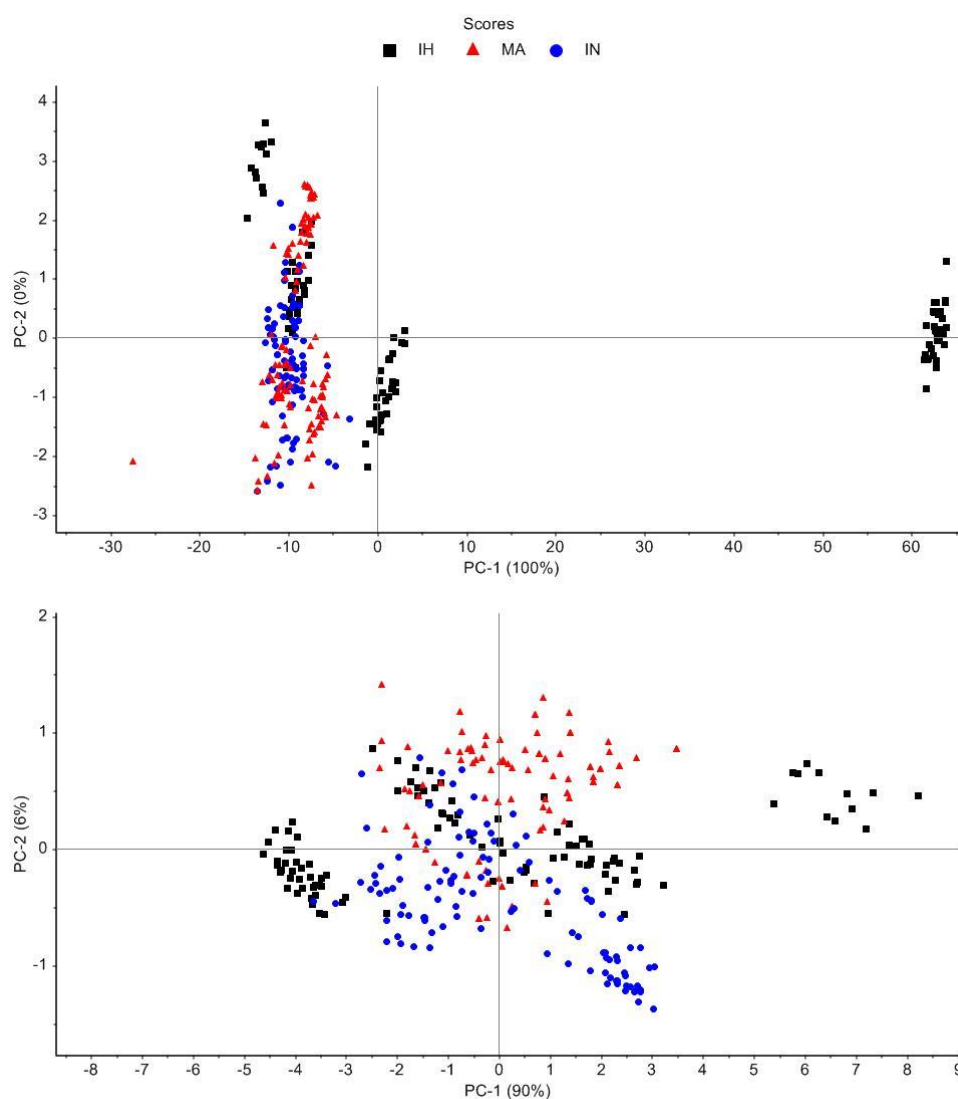
observados nesse estudo (Figura 1b), sendo assim, compreende-se que estas substâncias são relevantes para discriminar as espécies estudadas.

### 3.2. Análise de componente principal (PCA)

Na análise de PCA, foi possível notar um distanciamento de um grupo de amostras de *I. hederifolia* em relação aos demais quando os espectros NIR não foram pré-processados (Figura 2a). Este grupo de amostras foi o mesmo que ocasionou o desvio médio do espectro NIR desta espécie na Figura 1a. Apesar deste grupo de amostras ter se separado dos demais, não foi possível distinguir com clareza as três espécies. Além disso, a PCA do espectro NIR inteiro ( $4.0000-10.000\text{ cm}^{-1}$ ) sem nenhum pré-processamento não foi adequada, pois apenas um componente (PC1) representou 100% da variabilidade do conjunto amostral, inviabilizando sua aplicação para explicar diferenças entre as espécies.

Assim, na tentativa de melhorar os resultados da PCA, utilizou-se faixas do espectro NIR pré-processado. A Figura 2b ilustra a PCA com a seleção da faixa de  $4.240-5.540\text{ cm}^{-1}$  com o pré-processamento SNV. Nota-se que após a seleção da faixa e pré-processamento a variabilidade dos dados é melhor explicada pelos componentes (PC1=90% e PC2=6%) e para *I. hederifolia* ainda foi possível observar claramente a formação de subgrupos dentro do seu conjunto amostral. A justificativa para os subgrupos observados é a mesma relacionada ao deslocamento dos espectros (Figura 1a). Para esta espécie pequenas variações na espessura da folha, resultaram diferenças no espectro.

Contudo, apesar da aplicação dos pré-processamentos e seleção das faixas, não foi possível observar uma clara separação das espécies usando apenas uma análise não supervisionada (PCA). A grande sobreposição das amostras nas análises de PCA dificultou a identificação das espécies, o que corroborou os resultados obtidos dos espectros NIR (Figura 1). Assim, na tentativa de discriminar as espécies, foram utilizadas duas técnicas supervisionadas de classificação a PC-LDA e PLS-DA.



**Figura 2.** Valores de PC 1 e 2 para a análise de componente principal (PCA) para todo o espectro sem pré-processamento (a) e para a faixa de 4240-5540  $\text{cm}^{-1}$  com o pré-processamento SNV (b), para as espécies de *I. hederifolia* (quadrado preto), *I. nil* (triângulo vermelho) e *M. aegyptia* (círculo azul).

### 3.3. Análises discriminantes

Ao utilizar os espectros NIR com todas as faixas espectrais (4.000 a 10.000  $\text{cm}^{-1}$ ) apenas alguns pré-processamentos melhoraram a acurácia dos modelos em relação aos espectros brutos. Os modelos PC-LDA e PLS-DA sem o pré-processamento apresentaram valores de acurácia e índice de Youden superiores ou bem próximos aos com pré-processamento. Estes dados indicam que ao analisar

estes espectros o efeito dos pré-processamentos não levou a modelos mais acurados (Tabela 1).

Através da PC-LDA foi possível obter uma acurácia máxima de 96.19% para as espécies *I. hederifolia* e *M. aegyptia* e de 97,14% para a *I. nil* para o conjunto de calibração (Tabela 1). No conjunto de validação, os valores de acurácia maiores foram de 96,30%, 88,89% e 85,93% para discriminação das espécies *I. hederifolia*, *I. nil* e *M. aegyptia*, respectivamente (Tabela 1).

Na maioria dos casos, a acurácia do conjunto de dados de validação foi menor que no conjunto de dados de calibração. Essa é uma situação comum e significa que o conjunto de dados de validação possui um número suficiente de amostras (Borraz-Martínez et al., 2019). Além disso, esses resultados mostram como é importante validar os modelos gerados com amostras externas, que não foram usadas para o treinamento, para não incorrer no erro de selecionar modelos sobreajustados (do inglês: “overfitting”) (Pasquini, 2003).

Os pré-processamentos estudados não melhoraram a acurácia dos modelos em relação aos espectros sem o pré-processamento, em PCA-LDA, para as espécies *I. hederifolia* e *M. aegyptia* (Tabela 1). Este resultado pode ter sido observado em decorrência do desvio dos espectros que já existia nas amostras sem nenhum pré-processamento (Figura 1a). Já para a espécie *I. nil*, o pré-processamento SNV resultou em melhor acurácias dos modelos de calibração e de validação, sendo que, a acurácia de validação aumentou de 87,41% (sem pré-processamento) para 88.89% (com o pré-processamento SNV). Embora a discriminação das espécies por PC-LDA tenha atingido níveis de acurácia satisfatórios, ao considerar as figuras de mérito adotadas, a análise de PLS-DA foi a mais adequada para a classificação das três espécies.

O uso dos pré-processamentos usando a análise de PLS-DA também não ocasionou melhorias na acurácia de validação, com exceção dos pré-processamentos de 1SG e SNV+1SG para *I. hederifolia* e SNV para *I. nil*. Para *M. aegyptia* a melhor acurácia na validação foi no espectro sem o pré-processamento (94.81%) (Tabela 1). Contudo, para estudar as melhores faixas para discriminação para cada espécie, selecionou-se os dois melhores pré-processamentos considerando todas as figuras de mérito elegidas neste estudo.

**Tabela 1.** Resultados das análises discriminantes de PC-LDA e PLS-DA para as três espécies de corda-de-viola sem e com diferentes pré-processamentos na faixa espectral de 4.000-10.000 cm<sup>-1</sup>.

|                            | PC-LDA                 |                       |                    |                   |        | PLS-DA                 |                       |                    |                   |        |
|----------------------------|------------------------|-----------------------|--------------------|-------------------|--------|------------------------|-----------------------|--------------------|-------------------|--------|
| Pré-processamento          | Treinamento<br>(n=315) | Validação (n=135)     |                    |                   |        | Treinamento<br>(n=315) | Validação (n=135)     |                    |                   |        |
|                            | Ac.T (%) <sup>1</sup>  | Ac.V.(%) <sup>2</sup> | Sens. <sup>3</sup> | Esp. <sup>4</sup> | Youden | Ac.T (%) <sup>1</sup>  | Ac.V.(%) <sup>2</sup> | Sens. <sup>3</sup> | Esp. <sup>4</sup> | Youden |
| <i>Ipomoea hederifolia</i> |                        |                       |                    |                   |        |                        |                       |                    |                   |        |
| Sem pré-proces.            | 96,19                  | 96,30                 | 93,02              | 97,83             | 90,85  | 98,41                  | 95,56                 | 92,86              | 96,77             | 89,63  |
| SNV                        | 92,38                  | 85,93                 | 78,05              | 89,36             | 67,41  | 97,78                  | 93,33                 | 88,37              | 95,65             | 84,02  |
| Detrend                    | 90,48                  | 87,41                 | 96,30              | 85,19             | 81,48  | 98,73                  | 94,81                 | 85,71              | 100,00            | 85,71  |
| SNV+Detrend                | 93,33                  | 91,11                 | 91,67              | 90,91             | 82,58  | 98,10                  | 93,33                 | 85,11              | 97,73             | 82,83  |
| 1SG                        | 88,25                  | 75,56                 | 63,64              | 79,41             | 43,05  | 97,78                  | 97,78                 | 97,56              | 97,87             | 95,43  |
| SNV+1SG                    | 87,62                  | 78,52                 | 65,85              | 84,04             | 49,90  | 97,78                  | 96,30                 | 91,11              | 98,89             | 90,00  |
| <i>Ipomoea nil</i>         |                        |                       |                    |                   |        |                        |                       |                    |                   |        |
| Sem pré-proces.            | 96,51                  | 87,41                 | 88,64              | 86,81             | 75,45  | 96,83                  | 91,11                 | 89,80              | 91,86             | 81,66  |
| SNV                        | 97,14                  | 88,89                 | 90,91              | 87,91             | 78,82  | 96,51                  | 92,59                 | 88,68              | 95,12             | 83,80  |
| Detrend                    | 92,70                  | 74,81                 | 68,09              | 78,41             | 46,49  | 96,51                  | 86,67                 | 82,35              | 89,29             | 71,64  |
| SNV+Detrend                | 93,33                  | 77,78                 | 78,38              | 77,55             | 55,93  | 97,46                  | 88,89                 | 90,91              | 87,91             | 78,82  |
| 1SG                        | 85,08                  | 68,15                 | 59,09              | 72,53             | 31,62  | 94,60                  | 90,37                 | 86,54              | 92,77             | 79,31  |
| SNV+1SG                    | 85,71                  | 71,85                 | 66,67              | 73,96             | 40,63  | 94,29                  | 90,37                 | 89,58              | 90,80             | 80,39  |
| <i>Merremia aegyptia</i>   |                        |                       |                    |                   |        |                        |                       |                    |                   |        |
| Sem pré-proces.            | 96,19                  | 85,19                 | 95,83              | 82,88             | 78,72  | 97,14                  | 94,81                 | 92,68              | 95,74             | 88,43  |
| SNV                        | 90,16                  | 81,48                 | 86,96              | 80,36             | 67,31  | 97,14                  | 93,33                 | 86,67              | 96,67             | 83,33  |
| Detrend                    | 92,70                  | 85,93                 | 96,00              | 83,64             | 79,64  | 98,10                  | 90,37                 | 85,37              | 92,55             | 77,92  |
| SNV+Detrend                | 89,21                  | 83,70                 | 91,67              | 81,98             | 73,65  | 97,46                  | 94,07                 | 92,50              | 94,74             | 87,24  |
| 1SG                        | 88,57                  | 77,78                 | 68,75              | 80,58             | 49,33  | 95,87                  | 91,11                 | 89,47              | 91,75             | 81,23  |
| SNV+1SG                    | 83,81                  | 77,78                 | 68,75              | 80,58             | 49,33  | 96,19                  | 89,63                 | 85,00              | 91,58             | 76,58  |

(<sup>1</sup>Acurácia de treinamento; <sup>2</sup>Acurácia de validação; <sup>3</sup>Sensibilidade; <sup>4</sup>Especificidade).

O estudo de diferentes figuras de mérito é importante nesta etapa de seleção, em que se busca modelos mais robustos e acurados. Ao considerar a acurácia dos modelos, leva-se em conta o número de acertos totais, sejam eles verdadeiros positivos ou verdadeiros negativos; já o índice de Youden contribui na relação de equilíbrio entre a sensibilidade e a especificidade.

Porém, ao analisar o ponto de vista prático, no manejo de plantas daninhas, é desejável selecionar modelos que sejam capazes de identificar o alvo com melhor exatidão, ou seja, modelos com maior especificidade. Isso porque, quando uma planta não é identificada ela ficará sem controle e na safra seguinte e será responsável por gerar outras várias plantas, além de deixar sementes viáveis dormentes no solo (López-Granados, 2011; Piza, Nepomuceno e Alves, 2016; de Souza et al., 2020). Sendo assim, os pré-processamentos selecionados foram: 1SG e SNV+1SG para *I. hederifolia*; SNV e SNV+1SG para *I. nil*; e SNV e SNV+Detrend para *M. aegyptia*.

### 3.4. Seleção de faixas espectrais

Visando melhorar os resultados dos modelos obtidos, foi realizada a seleção de faixas espectrais, pois assim é possível identificar e excluir faixas com pouca informação sobre as espécies, melhorando a confiabilidade dos modelos (Borraz-Martínez et al., 2019). Para os espectros NIR sem pré-processamento, as faixas selecionadas de 4.240-7.460  $\text{cm}^{-1}$  e 6.800-7.460 + 4.240-5.540  $\text{cm}^{-1}$  apresentaram valores de acurácia de validação superiores aos observados para o espectro inteiro, para todas as espécies. Estas faixas representaram as regiões dos principais picos descritos na Figura 1.

Na faixa de 4.240-7.460  $\text{cm}^{-1}$  desconsiderou apenas as extremidades do espectro e na faixa de 6.800-7.460 + 4.240-5.540  $\text{cm}^{-1}$  desconsiderou as extremidades e o intervalo entre os dois principais picos (Figura 1b). Neste caso, a faixa com os melhores valores de acurácia foi a de 4.240-7.460  $\text{cm}^{-1}$ , ou seja, embora as maiores informações estejam nos picos, entre eles também há informações relevantes que podem ser usadas para melhorar a discriminação dessas espécies. Os valores de acurácia de validação referentes ao espectro todo (4.000 a 10.000  $\text{cm}^{-1}$ ) aumentaram de 95.56%, 91.11% e 94.81% para 98.52%, 94.07% e 96.30%, para *I. hederifolia*, *I. nil* e *M. aegyptia*, respectivamente, ao considerar na análise somente a faixa de 4.240-7.460  $\text{cm}^{-1}$  (Tabela 2).

Porém, ao considerar a especificidade nas faixas sem pré-processamento, não houve muita diferença entre as faixas de 4.240-7.460  $\text{cm}^{-1}$  e 6.800-7.460 + 4.240-5.540  $\text{cm}^{-1}$  (Tabela 2). Neste caso, a opção pela faixa que compreende apenas os dois picos, sem o intervalo entre eles, é uma opção melhor, em função do menor volume de dados. Em um estudo utilizando o espectro do visível e infravermelho foi observado que o uso de apenas quatro bandas apresentou bons resultados na diferenciação de espécies de plantas daninhas e cana-de-açúcar, simplificando, assim, os sensores ao utilizar um número menor de bandas espectrais para aplicação específica de herbicidas pós-emergentes (de Souza et al., 2020).

Ao contrário do comportamento observado para espectro NIR inteiro (4.000-10.000  $\text{cm}^{-1}$ ), os pré-processamentos nos espectros com as faixas selecionadas foram importantes para a otimização dos parâmetros dos modelos. Houve aumento nos valores das figuras de mérito, no estudo das faixas em relação ao espectro todo, com exceção da espécie *M. aegyptia* com o pré-processamento SNV, que obteve o melhor índice de Youden (83,33%) para análise do espectro todo (Tabela 2).

Considerando os valores de acurácia, especificidade e índice de Youden, as melhores faixas para todos os pré-processamentos ficaram entre 4.240-7.460 (dois picos + intervalo entre picos), 4.240-5.540 (primeiro pico) e 6.800-7.460 + 4.240-5.540  $\text{cm}^{-1}$  (dois picos – intervalo entre picos). Este resultado é otimista, pois sem aplicar os pré-processamentos só é possível atingir bons parâmetros nos modelos quando se considera o espectro inteiro ou grandes faixas do espectro. Neste estudo, a faixa de 4.240-5.540  $\text{cm}^{-1}$  (correspondente ao primeiro pico), possibilitou discriminar as três espécies utilizando um menor volume de dados, após a aplicação dos pré-processamentos.

Ao utilizar o pré-processamento de 1SG para discriminar *I. hederifolia* e SNV para *I. nil* e *M. aegyptia*, na faixa de 4.240-5.540  $\text{cm}^{-1}$ , foi possível obter excelentes resultados na acurácia de validação (99,26, 93,33 e 91,11%) e na especificidade (98,94, 95,18 e 97,65%), para cada espécie respectivamente (Tabela 2). Por outro lado, também foi possível notar que os piores resultados de discriminação foram observados na faixa de 6.800-7.460  $\text{cm}^{-1}$  (que compreende a área do segundo pico), com e sem os pré-processamentos.

**Tabela 2.** Resultados da análise de PLS-DA para três espécies de corda-de-viola sem e com diferentes pré-processamento em diferentes faixas espectrais.

| Faixa do espectro (cm <sup>-1</sup> ) | Train. <sup>1</sup> (n=315) | Validação (n=135)     |                    |                    |        | Train. (n=315)       | Validação (n=135)     |                    |                    |        | Train. (n=315)       | Validação (n=135)     |                    |                    |        |
|---------------------------------------|-----------------------------|-----------------------|--------------------|--------------------|--------|----------------------|-----------------------|--------------------|--------------------|--------|----------------------|-----------------------|--------------------|--------------------|--------|
|                                       | Ac. <sup>2</sup> (%)        | AcV. <sup>3</sup> (%) | Sens. <sup>4</sup> | Espe. <sup>5</sup> | Youden | Ac. <sup>2</sup> (%) | AcV. <sup>3</sup> (%) | Sens. <sup>4</sup> | Espe. <sup>5</sup> | Youden | Ac. <sup>2</sup> (%) | AcV. <sup>3</sup> (%) | Sens. <sup>4</sup> | Espe. <sup>5</sup> | Youden |
| <i>Ipomoea hederifolia</i>            |                             |                       |                    |                    |        |                      |                       |                    |                    |        |                      |                       |                    |                    |        |
|                                       | Sem pré-processamento       |                       |                    |                    |        | 1SG                  |                       |                    |                    |        | SNV+1SG              |                       |                    |                    |        |
| 4.000-10.000                          | 98,41                       | 95,56                 | 92,86              | 96,77              | 89,63  | 97,78                | 97,78                 | 97,56              | 97,87              | 95,43  | 97,78                | 96,30                 | 91,11              | 98,89              | 90,00  |
| 4.240-7.460                           | 98,73                       | 98,52                 | 97,62              | 98,92              | 96,54  | 100,00               | 98,52                 | 100,00             | 97,89              | 97,89  | 100,00               | 99,26                 | 100,00             | 98,94              | 98,94  |
| 6.800-7.460                           | 87,62                       | 86,67                 | 81,58              | 88,66              | 70,24  | 93,97                | 90,37                 | 83,72              | 93,48              | 77,20  | 95,24                | 90,37                 | 82,22              | 94,44              | 76,67  |
| 4.240-5.540                           | 99,05                       | 94,07                 | 92,50              | 94,74              | 87,24  | 100,00               | 99,26                 | 100,00             | 98,94              | 98,94  | 99,05                | 94,07                 | 92,50              | 94,74              | 87,24  |
| 6.800-7.460+<br>4.240-5.540           | 99,37                       | 97,78                 | 95,35              | 98,91              | 94,26  | 100,00               | 99,26                 | 100,00             | 98,94              | 98,94  | 99,37                | 97,78                 | 95,35              | 98,91              | 94,26  |
| <i>Ipomoea nil</i>                    |                             |                       |                    |                    |        |                      |                       |                    |                    |        |                      |                       |                    |                    |        |
|                                       | Sem pré-processamento       |                       |                    |                    |        | SNV                  |                       |                    |                    |        | SNV+1SG              |                       |                    |                    |        |
| 4.000-10.000                          | 96,83                       | 91,11                 | 89,80              | 91,86              | 81,66  | 96,51                | 92,59                 | 88,68              | 95,12              | 83,80  | 94,29                | 90,37                 | 89,58              | 90,80              | 80,39  |
| 4.240-7.460                           | 98,73                       | 94,07                 | 95,74              | 93,18              | 88,93  | 98,10                | 92,59                 | 91,84              | 93,02              | 84,86  | 100,00               | 98,52                 | 96,23              | 100,00             | 96,23  |
| 6.800-7.460                           | 78,73                       | 71,85                 | 63,83              | 76,14              | 39,97  | 89,84                | 86,67                 | 78,95              | 92,31              | 71,26  | 93,65                | 89,63                 | 83,64              | 93,75              | 77,39  |
| 4.240-5.540                           | 96,83                       | 88,15                 | 92,68              | 86,17              | 78,85  | 97,46                | 93,33                 | 90,38              | 95,18              | 85,57  | 100,00               | 97,78                 | 94,44              | 100,00             | 94,44  |
| 6.800-7.460+<br>4.240-5.540           | 97,14                       | 93,33                 | 92,00              | 94,12              | 86,12  | 98,10                | 92,59                 | 91,84              | 93,02              | 84,86  | 100,00               | 96,30                 | 94,23              | 97,59              | 91,82  |
| <i>Merremia aegyptia</i>              |                             |                       |                    |                    |        |                      |                       |                    |                    |        |                      |                       |                    |                    |        |
|                                       | Sem pré-processamento       |                       |                    |                    |        | SNV                  |                       |                    |                    |        | SNV+Detrend          |                       |                    |                    |        |
| 4.000-10.000                          | 97,14                       | 94,81                 | 97,46              | 97,46              | 97,46  | 97,14                | 93,33                 | 86,67              | 96,67              | 83,33  | 97,46                | 97,46                 | 92,50              | 94,74              | 87,24  |
| 4.240-7.460                           | 99,68                       | 96,30                 | 98,10              | 98,10              | 98,10  | 98,10                | 90,37                 | 80,85              | 95,45              | 76,31  | 98,10                | 98,10                 | 90,48              | 95,70              | 86,18  |
| 6.800-7.460                           | 87,30                       | 84,44                 | 96,83              | 96,83              | 96,83  | 86,67                | 87,41                 | 93,10              | 85,85              | 78,95  | 96,83                | 96,83                 | 82,61              | 95,51              | 78,11  |
| 4.240-5.540                           | 98,73                       | 92,59                 | 98,10              | 98,10              | 98,10  | 97,14                | 91,11                 | 80,00              | 97,65              | 77,65  | 98,10                | 98,10                 | 92,68              | 95,74              | 88,43  |
| 6.800-7.460+<br>4.240-5.540           | 99,37                       | 94,81                 | 98,73              | 98,73              | 98,73  | 97,46                | 91,11                 | 81,25              | 96,55              | 77,80  | 98,73                | 98,73                 | 97,44              | 95,83              | 93,27  |

(<sup>1</sup>Treinamento; <sup>2</sup>Acurácia de Treinamento; <sup>3</sup>Acurácia de validação; <sup>4</sup>Sensibilidade; <sup>5</sup>Especificidade).

Embora os dois picos somados tenham apresentado bons resultados, quando analisados separadamente, o primeiro pico do espectro (Figura 1b), que corresponde a faixa de 4.240-5.540  $\text{cm}^{-1}$ , possui um melhor desempenho na discriminação das espécies em relação ao segundo pico (6.800-7.460  $\text{cm}^{-1}$ ).

Sendo assim, o estudo contribuiu para identificar as faixas do espectro que contém as maiores informações para discriminar as espécies, assim como os melhores pré-processamentos que, conseqüentemente, poderão ajudar no desenvolvimento de sensores a bordo de máquinas menos complexos, que necessitem de um número menor informações para conseguir bons resultados de discriminação das espécies de corda-de-viola.

Os trabalhos futuros devem se concentrar em entender a influência de outros fatores no comportamento espectral dessas espécies, assim como, estágios de desenvolvimento, variações nutricionais e geográficas.

#### 4. CONCLUSÃO

Conclui-se que a espectroscopia NIR associada às análises de PC-LDA e PLS-DA pode ser utilizada para discriminar espécies de corda-de-viola. O estudo das faixas espectrais bem como os pré-processamentos foram relevantes e promoveram melhorias na capacidade de acerto dos modelos. A acurácia de validação para os modelos gerados com todo o espectro (4.000 a 10.000  $\text{cm}^{-1}$ ) foi de até 97,78, 92,59 e 94,07%, e após estudo de faixas e pré-processamentos esses valores foram de até 99,26, 98,52 e 98,73%, para *I. hederifolia*, *I. nil* e *M. aegyptia*, respectivamente.

## 5. REFERÊNCIAS

- 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):1–9.
- Braga AF, Barroso AAM, Amaral CL, Nepomuceno MP, Alves PLCA (2018) Population interference of glyphosate resistant and susceptible ryegrass on eucalyptus initial development. **Planta Daninha** 36:1-10.
- Carvalho S, Schlerf M, van der Putten WH, Skidmore AK (2013) Hyperspectral reflectance of leaves and flowers of an outbreak species discriminates season and successional stage of vegetation. **International Journal of Applied Earth Observation and Geoinformation** 24(1):32–41.
- 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.
- Cunha Júnior LC, Teixeira GHA, Nardini V, Walsh KB (2016) Quality evaluation of intact açaí and juçara fruit by means of near infrared spectroscopy. **Postharvest Biology and Technology** 112:64–74.
- Dammer KH, Wartenberg G (2007) Sensor-based weed detection and application of variable herbicide rates in real time. **Crop Protection** 26(3):270–277.
- De Souza MF, Amaral LR, Oliveira SRM, Coutinho MAN, Ferreira Netto C (2020) Spectral differentiation of sugarcane from weeds. **Biosystems Engineering** 190:41–46.
- De Souza Rolim G, De Camargo MBP, Grosselilania D, De Moraes JFL (2007) Climatic classification of köppen and thornthwaite systems and their applicability in the determination of agroclimatic zoning for the state of São Paulo, Brazil. **Bragantia** 66(4):711–720.
- Hoagland D, Arnon D (1938) The water-culture method for growing plants without soil. **Agricultural Experiment Station Circular 347** (University of California, Berkeley).
- Jamshidi B, Mohajerani E, Jamshidi J (2016) Developing a Vis/NIR spectroscopic system for fast and non-destructive pesticide residue monitoring in agricultural product. Measurement: **Journal of the International Measurement Confederation** 89:1–6.
- Jenett-Siems K, Schimming T, Kaloga M, Eich E, Siems K, Gupta MP, Witte L, Hartmann T (1998) Pyrrolizidine alkaloids of *Ipomoea hederifolia* and related species. **Phytochemistry**, 47(8):1551–1560.
- Jintao X, Yufei L, Liming Y, Chunyan L, Quanwei Y, Weiyang 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.

Kennard ARW, Stone LA (1969) Computer aided design of experiments. **American Society for Quality** 11(1):137–148.

López-Granados F (2011) Weed detection for site-specific weed management: Mapping and real-time approaches. **Weed Research** 51(1):1–11.

Meira M, da Silva EP, David JM, David JP (2012) Review of the genus Ipomoea: Traditional uses, chemistry and biological activities. **Brazilian Journal of Pharmacognosy** 22(3):682–713.

Nicolai BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J (2007) Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. **Postharvest Biology and Technology** 46(2):99–118.

Pasquini C (2003) Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society** 14(2):198–219.

Piza CST, Nepomuceno MP, Alves PLDCA (2016) Period prior to interference of morning glory in sugarcane. **Científica** 44(4):543.

Rodrigues BN, Almeida FS (2018) **Guia de herbicidas**. 7.ed. Curitiba: Produção Independente, v.1.

Snel FA, Braga JWB, da Silva D, Wiedenhoeft AC, Costa A, Soares R, Coradin VTR, Pastore T C. M. (2018) Potential field-deployable NIRS identification of seven Dalbergia species listed by CITES. **Wood Science and Technology** 52(5):1411–1427.

Soares LF, Da Silva DC, Bergo MCJ, Coradin VTR, Braga JWB, Pastore TCM (2017) Avaliação de espectrômetro nir portátil e pls-da para a discriminação de seis espécies similares de madeiras Amazônicas. **Química Nova** 40(4):418–426.

Steidle Neto AJ, Lopes DC, Toledo JV, Zolnier S, Silva TGF (2018) Classification of sugarcane varieties using visible/near infrared spectral reflectance of stalks and multivariate methods. **Journal of Agricultural Science** 156(4):537–546.

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(5):277–284.

Xiaobo Z, Jiewen Z, Povey MJW, Holmes M, Hanpin M (2010) Variables selection methods in near-infrared spectroscopy. **Analytica Chimica Acta** 667(1–2):14–32.

Youden WJ (1950) Index for rating diagnostic tests. **Cancer** 3(1):32–35.

### CAPÍTULO 3 – Espectroscopia NIR e métodos multivariados para discriminação de espécies de cordas-de-viola em diferentes estádios de crescimento

**RESUMO** - Espécies de corda-de-viola são plantas daninhas muito comuns em cultivos tropicais e ocasionam prejuízos diretos e indiretos. O manejo destas espécies é realizado principalmente pelo uso de herbicidas, em aplicações em pré e pós-emergência em área total. Estudos abordando novas técnicas para identificação dessas espécies poderão contribuir para o desenvolvimento de sensores proximais para realização de um manejo direcionado no momento ideal. Desta forma, o objetivo deste trabalho foi utilizar a espectroscopia no infravermelho próximo (NIRS) e análises multivariadas para discriminar duas espécies de corda-de-viola em três estádios de crescimento. Os espectros NIR foram coletados nas folhas de *Ipomoea hederifolia* nos estádios 10, 11 e 12 e nas folhas de *Merremia aegyptia* nos estádios 11, 12 e 14, na faixa espectral de 4.000 a 10.000  $\text{cm}^{-1}$ . Utilizou-se a PCA e PC-LDA para compreensão do espectro todo e faixas específicas. A espectroscopia NIR associada a PCA e PC-LDA foram suficientes para discriminar as espécies de *I. hederifolia* e *M. aegyptia* e seus estádios de crescimento. A PCA permitiu uma boa segregação dos estádios quando aplicada individualmente nas espécies e das espécies quando aplicada individualmente nos dois primeiros estádios. A PC-LDA classificou corretamente entre 90,93 a 100% as espécies e os estádios. Os melhores resultados de discriminação foram observados nas faixas de 4.500 a 6.000  $\text{cm}^{-1}$  e 4.500 a 6.000 + 6.500 a 7.750  $\text{cm}^{-1}$ .

**Palavras-chave:** agricultura de precisão, assinatura espectral, *Ipomoea hederifolia*, manejo específico, *Merremia aegyptia*, plantas daninhas

**ABSTRACT** – Morning glory species are very common weeds in tropical crops, where they cause direct and indirect damage. The management of these species is carried out mainly through the use of herbicides in pre and post-emergence applications. Studies that address new techniques for the identification of these species can contribute to the development of proximal sensors to perform specific and rational management. Thus, the objective of this work was to use near-infrared spectroscopy (NIRS) and multivariate analyzes to discriminate two species of morning glory in three stages of growth. The NIR spectra were collected in the leaves of *Ipomoea hederifolia* in stages 10, 11, and 12 and in the leaves of *Merremia aegyptia* in stages 11, 12, and 14, in the spectral range of 4,000 to 10,000  $\text{cm}^{-1}$ . PCA and PC-LDA were used to understand the entire spectrum and specific ranges. NIRS associated with PCA and PC-LDA were sufficient to discriminate the species of *I. hederifolia* and *M. aegyptia* and their growth stages. PCA allowed good segregation of the stages when applied individually to the species and when applied individually in the first and two-stage. PC-LDA correctly classified between 90.93 and 100% of the species and stages. The best discrimination results were observed in the ranges of 4,500 to 6,000  $\text{cm}^{-1}$  and 4,500 to 6,000 + 6,500 to 7,750  $\text{cm}^{-1}$ .

**Keywords:** *Ipomoea hederifolia*, *Merremia aegyptia*, precision agriculture, specific management, spectral signature, weeds

## 1. INTRODUÇÃO

A agricultura tem passado por uma grande revolução tecnológica. Sensores, robôs, algoritmos e inteligência artificial são palavras cada vez mais presentes no campo (Klerkx et al., 2019). Neste contexto, o principal objetivo de se incluir novas tecnologias em sistemas agrícolas está relacionado à adoção de técnicas de manejo que possibilitem o aumento de produtividade de maneira sustentável (Van Evert et al., 2017).

O manejo de plantas daninhas é fundamental para alcançar altas produtividades e comumente é realizado pelo uso de herbicidas, aplicados em área total (Piza et al., 2016; Alves et al., 2018). A redução no uso de insumos é uma das principais metas no avanço tecnológico, por impactarem diretamente no custo operacional e na redução de possíveis danos ao meio-ambiente (Klerkx et al., 2019). Por isso, com o apressar das tecnologias, estudos têm apresentado várias alternativas para realização de aplicações de defensivos somente na planta-alvo, reduzindo significativamente a quantidade dos insumos aplicados, sem a perda de eficiência, mas com menor impacto ambiental (Wang et al., 2007; Hansen et al., 2013; Singh et al., 2020).

Algumas destas alternativas para otimização do uso de herbicidas é a identificação precisa do alvo, ou seja, de plantas daninhas, por meio de sensores, imagens e dados espectrais (Wang et al., 2007; Hansen et al., 2013; Souza et al., 2020). A vantagem do desenvolvimento de sensores frente ao uso de outras tecnologias está relacionada ao tempo de processamento das informações, interferindo diretamente na tomada de decisão e efetividade de controle (López-Granados et al., 2010; Souza et al., 2020).

Já existem sensores de plantas daninhas disponíveis comercialmente, mas estes só detectam com eficiência a presença ou a ausência dessas espécies em contraste com o solo ou a palhada (Wang et al., 2007, Lopez-Granados et al., 2010; Singh et al., 2020). O desenvolvimento de sensores para discriminação entre espécies de plantas daninhas é possível, porém demanda a seleção de melhores bandas e estudos de calibração, que envolvam a compreensão das propriedades espectrais em relação a outros fatores fisiológicos, como o teor de clorofilas, o teor de água e a presença ou ausência de estresse biótico e abiótico (Souza et al., 2020).

A família das Convolvulaceae compreende espécies de plantas daninhas bastante expressivas, capazes de ocasionar prejuízos diretos e indiretos às culturas (Piza et al., 2016; Alves et al., 2018; Ribeiro et al., 2018). Os prejuízos diretos estão relacionados à competição por nutrientes, luz e água, que chegam a reduzir em até 36% a produtividade (Bhullar et al., 2012; Pagnoncelli et al., 2017). Já os prejuízos indiretos estão relacionados à colheitabilidade; por serem espécies de hábito trepador, seus ramos se entrelaçam nas plantas comprometendo o rendimento operacional da colhedora e, consequentemente, impactando nos custos operacionais (Piza et al., 2016; Alves et al., 2018).

A aplicação de técnicas quimiométricas juntamente com a espectroscopia no infravermelho próximo (NIRS), contribuem na elaboração de modelos de discriminação de plantas daninhas e são eficazes na discriminação de espécies da mesma família e do mesmo gênero (Capítulo 2). Porém, compreende-se que a identificação dos estádios de crescimento é tão importante quanto discriminar as espécies, visto que o estágio interfere na suscetibilidade das espécies aos herbicidas (Rodrigues e Almeida, 2015; Ribeiro et al., 2018).

Em virtude do relatado, o objetivo desse trabalho foi gerar um modelo para discriminar duas espécies de corda-de-viola e seus estádios de crescimento.

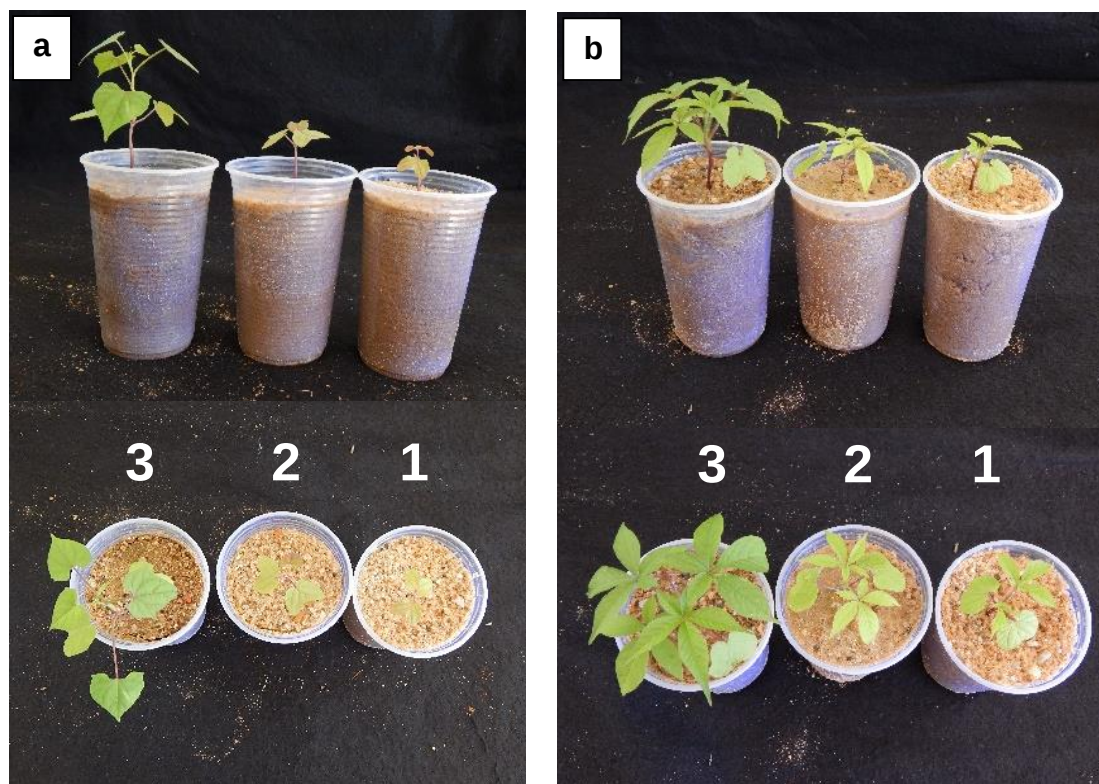
## 2. MATERIAL E MÉTODOS

### 2.1. Material Vegetal

As espécies de plantas daninhas selecionadas foram *Ipomoea hederifolia* e *Merremia aegyptia*, todas herbáceas, de hábito trepador e pertencentes à família Convolvulaceae (Figura 1).

A semeadura destas espécies foi realizada de forma escalonada, em três épocas, em vasos plásticos com capacidade para 0,5 litros, preenchidos com areia de rio lavada e peneirada. Após a emergência das espécies, realizou-se o desbaste, deixando apenas uma planta por vaso. Todas as plantas foram cultivadas a céu aberto, sob coordenadas 21° 14' 38.6" Sul e 48° 17' 56.3" Oeste. Foram cultivados 50 vasos para cada espécie em cada um dos estádios de desenvolvimento, totalizando 150 vasos por espécie. Os vasos foram irrigados diariamente com água e solução nutritiva completa de Hoagland e Arnon (1938).

Na primeira semana, a solução nutritiva foi ajustada para uma concentração de 25%, na segunda semana para 50% e assim consecutivamente até que atingisse 100% de força iônica. O período de condução das plantas foi entre os meses de março a abril de 2019, e o clima da região, segundo a classificação de Köppen, é Aw.



**Figura 1.** *Ipomoea hederifolia* (a) e *Merremia aegyptia* (b) nos estádios de crescimento 1, 2 e 3.

No dia da coleta dos espectros, as plantas foram classificadas quanto ao estágio de desenvolvimento, conforme o BBCH-code (Bleholder et al., 1991). Todas espécies estavam no macro-estádio 1 (desenvolvimento de folhas), porém, devido às particularidades de desenvolvimento de cada espécie, os micro-estádios foram os seguintes: 10, 11 e 12 para *I. hederifolia* e 11, 12 e 14 para *M. aegyptia* (Tabela 1 e Figura 1).

**Tabela 1.** Código decimal para a classificação dos estádios de desenvolvimento de plantas

| BBCH-code | Definição                                                 |
|-----------|-----------------------------------------------------------|
| 1         | Desenvolvimento de folhas (gema principal)                |
| 10        | Folhas cotiledonares completamente desdobradas;           |
| 11        | 1ª folha/1º par de folhas (ou folha composta) desdobrada; |
| 12        | 2ª folha/2º par de folhas (ou folha composta) desdobrada; |
| 13        | 3ª folha/3º par de folhas (ou folha composta) desdobrada; |
| 14        | 4ª folha/4º par de folhas (ou folha composta) desdobrada; |

Fonte: Adaptado de Bleiholder et al. (1991).

Para fins caracterizar os estádios e as espécies, foi calculado o teor de água nas folhas de cada espécie em cada um dos micro-estádios. Para essa etapa foram amostradas 15 plantas de cada estágio e espécie e determinou-se a massa fresca de folhas. Na sequência, estas folhas foram colocadas para secar em estufa com circulação forçada de ar a 40°C por 96 horas, após as quais determinou-se a massa seca. O teor de água nas folhas foi expresso em percentagem, sendo calculado através da seguinte equação:

$$TA(\%) = \left( \frac{x - y}{x} \right) \times 100$$

Onde  $x$ , representa a massa fresca (g) das folhas e  $y$  representa a massa seca (g) das folhas. Após o cálculo do teor de água, os dados foram submetidos a ANOVA {fatorial de 3 (estádios) x 2 (espécies) em delineamento inteiramente casualizado} e as médias foram comparadas pelo teste de Tukey a 5% de probabilidade.

## 2.2. Obtenção dos espectros

Os espectros NIR foram obtidos através do acessório NIR Fiber Optic Probe acoplado ao espectrofotômetro FT-NIR Spectrum 100N (Perkin Elmer, Shelton, Estados Unidos da América). Os espectros de refletância difusa foram adquiridos na forma de  $\log(1/R)$ , onde  $R$  correspondeu à refletância, na faixa espectral de 10.000 a 4.000  $\text{cm}^{-1}$  (1.000 a 2.500 nm), com 32 scans, a uma resolução espectral de 4  $\text{cm}^{-1}$  e com intervalo de 2  $\text{cm}^{-1}$ . No momento da coleta dos espectros NIR, a temperatura do ambiente foi mantida a 23 °C  $\pm$  1, permanecendo constante durante todo o processo de aquisição espectral.

A leitura dos espectros foi realizada na parte adaxial das folhas frescas, posicionadas sobre um fundo preto (de E.V.A). Adotou-se o uso do fundo preto para reduzir a interferência da entrada de luz externa. Para cada estágio foram realizadas 100 leituras do espectro ( $n=300$ , para cada espécie).

### 2.3. Quimiometria

Os dados espectrais foram analisados usando o programa Unscrambler®, versão 10.3 (CAMO. Oslo. Noruega). O banco de dados espectral foi dividido em treinamento (70%) e validação (30%) através do algoritmo de Kennard-Stone (Kennard e Stone, 1969). O pré-processamento usado nos espectros foi o “Standard Normal Variate” (SNV) somado a 1ª derivada de Savitzky–Golay (SNV+1SG) agrupado em 6 e 6 pontos, conforme definido em estudos anteriores (Braga et al., 2021, capítulo 2).

A análise de componente principal (PCA) foi realizada com todo o espectro NIR ( $4.000$  a  $10.000\text{ cm}^{-1}$ ) e nas faixas de  $4.500$  a  $6.000\text{ cm}^{-1}$  e  $6.500$  a  $7.750\text{ cm}^{-1}$ . As faixas espectrais foram selecionadas conforme os valores dos loadings da PCA e os picos de maior refletância, observados nos espectros. Para discriminar os estádios das espécies foram desenvolvidos modelos de classificação usando a análise discriminante linear de componentes principais (PC-LDA), para todos os estádios e espécies juntos (6 classificadores) com e sem o pré-processamento de SNV+1SG.

O método de separação dentro da PC-LDA foi pela distância de Mahalanobis com o uso de 10 componentes principais (PCs) e o método de validação cruzada aleatória, com 10 amostras por segmento. O desempenho dos classificadores foi avaliado pela percentagem de verdadeiros positivos.

## 3. RESULTADOS E DISCUSSÃO

### 3.1. Caracterização do material vegetal e espectros NIR

Os valores médios dos teores de umidade em *I. hederifolia* nos estádios 1, 2 e 3 (correspondentes aos micro-estádios: 10, 11 e 12) e *M. aegyptia* nos

estádios 1, 2 e 3 (correspondentes aos micro-estádios: 11, 12 e 14) foram significativamente diferentes para as espécies e estádios (Tabela 2). Houve diferença significativa ( $p < 0,05$ ) apenas para os fatores isolados, sendo que a interação entre os fatores espécies e estádios não foi significativa ( $p = 0,8073$ ). *Ipomoea hederifolia* apresentou maior teor de umidade (82,89%) que *M. aegyptia* (80,78%). Já para a comparação entre os estádios, a maior diferença observada no teor de umidade foi entre o estágio 1 (78,93%), comparado aos estádios 2 (83,33%) e 3 (83,24%). Embora ambas espécies estivessem nas mesmas condições de manejo, essas variações eram esperadas em função da turgidez dos tecidos serem características de cada espécie e estágio (Barroso et al., 2017).

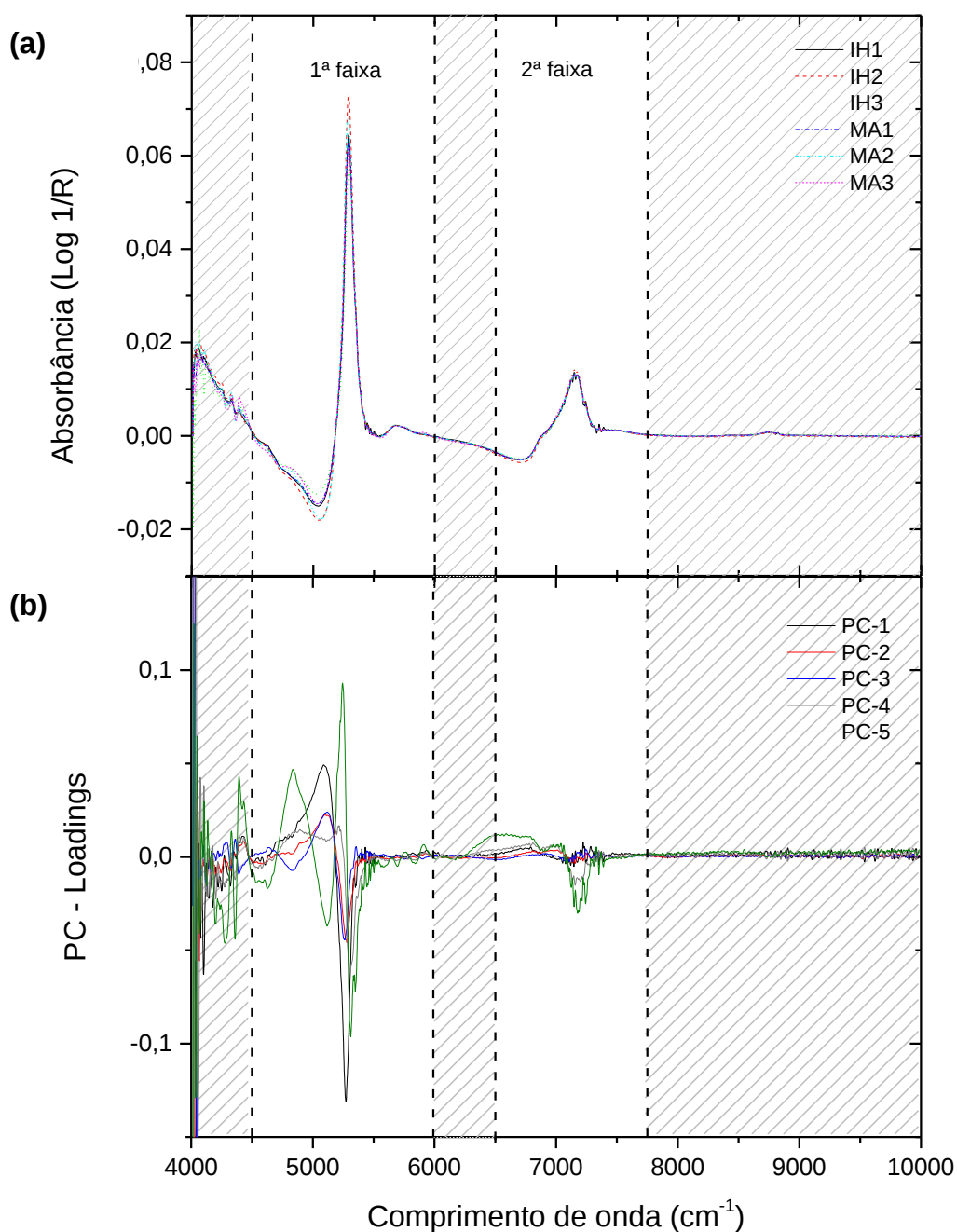
**Tabela 2.** Valores médios do teor de água para *I. hederifolia* e *M. aegyptia* nos três estádios de crescimento.

|                       | <b>Estádio 1</b> | <b>Estádio 2</b> | <b>Estádio 3</b> | <b>Médias*</b> |
|-----------------------|------------------|------------------|------------------|----------------|
| <i>I. hederifolia</i> | 80,33            | 84,25            | 84,09            | 82,89 a        |
| <i>M. aegyptia</i>    | 77,53            | 82,41            | 82,40            | 80,78 b        |
| <b>Médias*</b>        | 78,93 B          | 83,33 A          | 83,24 A          | <b>81,83</b>   |

\* Médias seguidas da mesma letra, maiúsculas na linha e minúsculas na coluna, não diferem entre si, ao nível de 5% de probabilidade.

Os espectros NIR médios das espécies nos três estádios na faixa de 4.000 a 10.000  $\text{cm}^{-1}$  com o pré-processamento SNV+1SG apresentaram dois principais picos de absorbância (Figura 2 a). As regiões destes picos coincidem com os picos de absorbância da água, que são próximos de 6.800 e 5.500  $\text{cm}^{-1}$ , indicando os altos teores de umidade nas amostras de corda-de-viola (Cunha Júnior et al., 2015; Torres et al., 2019).

A faixa de absorção da água é bastante característica no espectro do infravermelho, em função das vibrações de OH serem facilmente detectadas (Pasquini, 2003; Wang et al., 2015). Uma das primeiras aplicações da espectroscopia de refletância difusa na faixa do infravermelho foram justamente com a finalidade de explorar métodos alternativos para determinação da umidade de maneira não destrutiva (Pasquini, 2003; Torres et al., 2019).



**Figura 2.** Absorbância média dos espectros (a) e PC-loadings (b) de 4.000 a 10.000  $\text{cm}^{-1}$  com o pré-processamento SNV+1SG das espécies de *I. hederifolia* nos estádios 10 (IH1); 11 (IH2) e 12 (IH3); e *M. aegyptia* nos estádios 11 (MA1); 12 (MA2) e 14 (MA3). (1ª faixa: 4.500 a 6.000  $\text{cm}^{-1}$ ; 2ª faixa: 6.500 a 7.750  $\text{cm}^{-1}$ ).

No entanto, embora as vibrações de OH expliquem em maioria os elevados picos de absorbância, nessas faixas também são identificados outros compostos orgânicos, como flavonoides, alcaloides e antocianinas (León e Downey, 2006; Jintão et al., 2018; Amanah et al., 2020). Pasquini (2003) ressalta que as regiões com maior absorbância são as que contém as variações mais

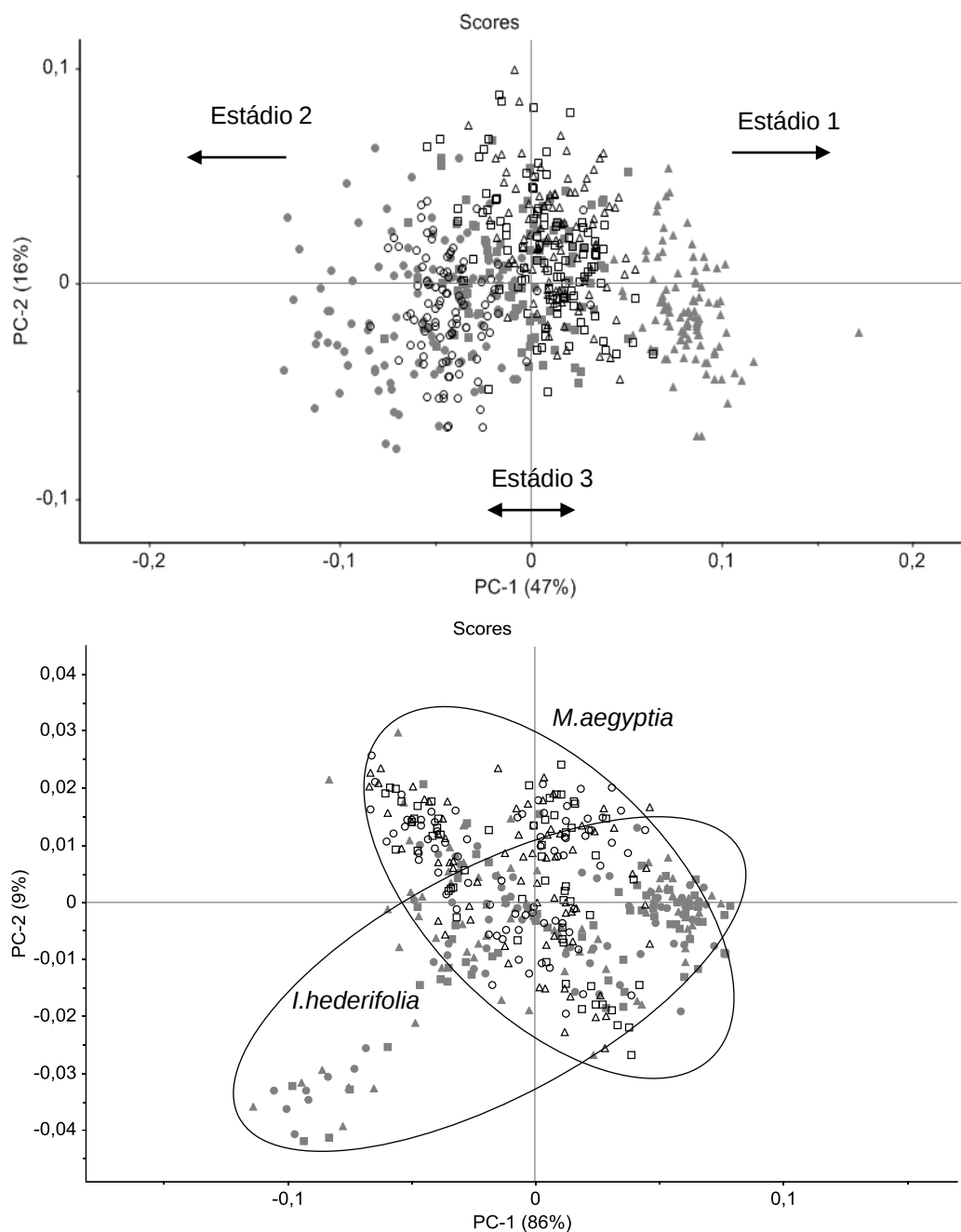
consideráveis para compreensão do comportamento espectral das amostras. Sendo assim, afim de melhorar a compreensão das áreas de maior influência, selecionou-se as faixas com os maiores picos e com os maiores valores de “loadings” (Figura 2 b) para o estudo das análises multivariadas.

### 3.2. Análise de componente principal (PCA)

A análise dos componentes principais (PCA) do espectro todo (4.000 a 10.000  $\text{cm}^{-1}$ ) com o pré-processamento de SNV+1SG explicou 47% das variáveis em PC1 e 16% em PC2, ou seja, 63% nos dois primeiros componentes (Figura 3a). Houve uma segregação razoável dos grupos na PCA, mas ao centro do hiperplano as amostras ainda ficaram sobrepostas. As amostras de *I. hederifolia* segregaram o segundo e o primeiro estágio, sendo que as amostras do segundo estágio se associaram a valores negativos de PC1 e do primeiro, a valores positivos. Já as amostras do terceiro estágio ficaram ao centro, com valores positivos e negativos para PC1 e PC2.

A espécie *M. aegyptia* segregou menos, mas apresentou o mesmo comportamento de *I. hederifolia*, mantendo a maioria das amostras do estágio 2 no quadrante negativo de PC1 e do estágio 1 no quadrante positivo. As setas no gráfico indicam a orientação das amostras de cada estágio (Figura 3 a). Portanto, a PCA aplicada ao espectro todo (4.500 a 10.000  $\text{cm}^{-1}$ ) não foi suficiente para discriminar as espécies e estádios.

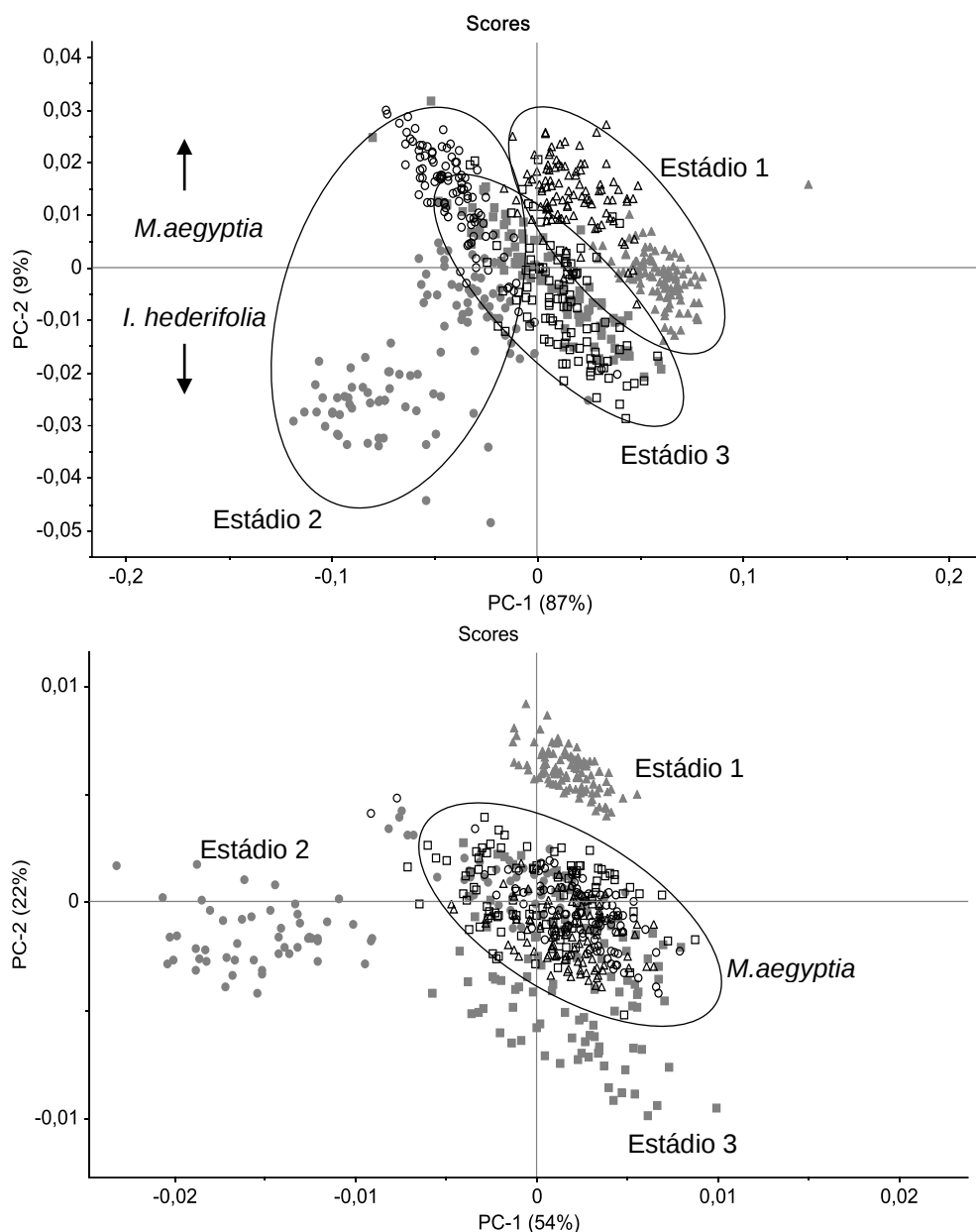
Ao considerar somente as duas faixas (4.500 a 6.000  $\text{cm}^{-1}$  + 6.500 a 7.750  $\text{cm}^{-1}$ ) (Figura 3 b), a soma dos dois primeiros componentes foi mais representativa para o estudo da variabilidade (PC1 = 86% e PC2 = 9%), mas as amostras também não segregaram. Houve muita sobreposição entre as espécies e não foi possível identificar grupos de separação dos estádios.



**Figura 3.** Valores de PC 1 e 2 resultantes da análise de componente principal (PCA) com o espectro todo (a) e com as duas faixas (b) para as espécies *I. hederifolia* (cinza) e *M. aegyptia* (preto) em três estádios de desenvolvimento. (triângulo: estágio 1; círculo: estágio 2; quadrado: estágio 3).

A PCA na 1ª faixa, de 4.500 a 6.000  $\text{cm}^{-1}$  (Figura 4 a), discriminou os estádios com pouca sobreposição e também as espécies dentro de cada estágio, com exceção para o estágio 3, em que as amostras das duas espécies ficaram bastante sobrepostas. PC1 e PC2 explicaram 87% e 9% da variabilidade

espectral total, respectivamente. Estes resultados foram melhores do que os observados na PCA com o espectro todo (Figura 3 a). Ao analisar somente essa região do espectro NIR, foi possível notar diferenças não somente entre os estádios, mas também entre as espécies. As amostras de *M. aegyptia* dos estádios 1 e 2 se apresentaram em maioria do quadrante positivo de PC2 e *I. hederifolia* apresentou maioria do quadrante negativo de PC2.



**Figura 4.** Valores de PC 1 e 2 resultantes da análise de componente principal (PCA) com a 1ª faixa (a) e com a 2ª faixa (b) para as espécies de *I. hederifolia* (cinza) e *M. aegyptia* (preto) em três estádios de desenvolvimento. (triângulo: estágio 1; círculo: estágio 2; quadrado: estágio 3).

Na 2ª faixa, de 6.500 a 7.750  $\text{cm}^{-1}$  (Figura 4 b), PC1(54%) e PC2(22%) foram menos representativos que a 1ª faixa e segregaram somente algumas amostras de *I. hederifolia* e seus estádios. *Merremia aegyptia*, nessa faixa, permaneceu ao centro do hiperplano, com uma alta sobreposição dos estádios. Sendo assim, compreende-se que para o estudo dos estádios e espécies a 1ª faixa (4.500 a 6.000  $\text{cm}^{-1}$ ) é a mais adequada.

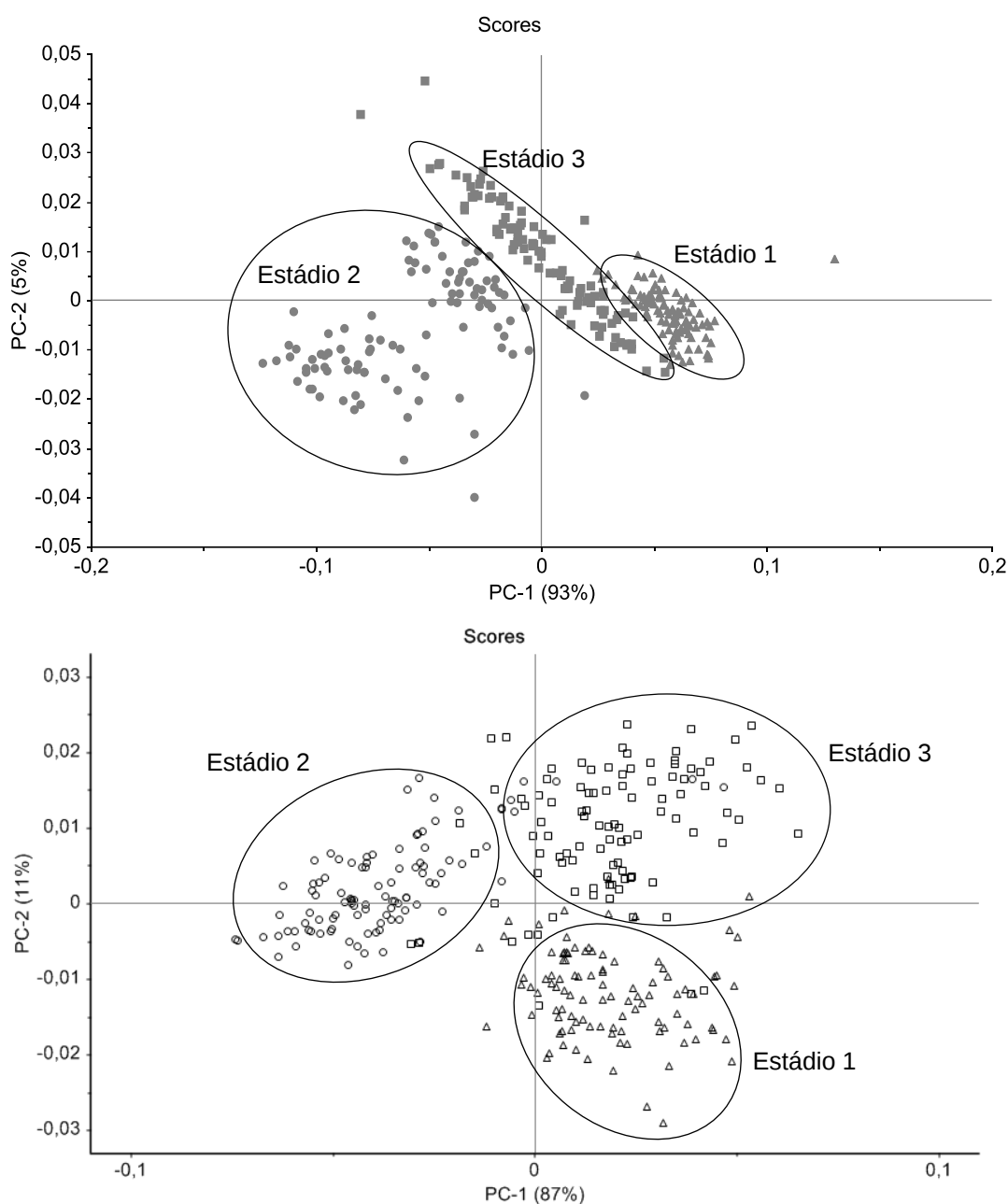
A PCA para cada espécie na 1ª faixa, de 4.500 a 6.000  $\text{cm}^{-1}$ , discriminou muito bem os estádios de *I. hederifolia* (Figura 5 a) e *M. aegyptia* (Figura 5 b) e mantiveram uma influência alta dos PC1 e PC2. A variabilidade foi explicada em 93% e 5% para *I. hederifolia*, e em 87% e 11%, para *M. aegyptia*, em PC1 e PC2, respectivamente. Os estádios em *I. hederifolia* se separaram da seguinte forma: estágio 1, no quadrante positivo de PC1; estágio 2, no quadrante negativo de PC1 e estágio 3, no quadrante positivo de PC2. Já para os estádios de *M. aegyptia*, a separação das amostras ficou ainda mais evidente: estágio 1, quadrante positivo de PC1 e negativo de PC2; estágio 2, quadrante negativo de PC1 e estágio 3, quadrantes positivos de PC1 e PC2.

A PCA aplicada a cada uma das espécies separadas na 2ª faixa (6.500 a 7.750  $\text{cm}^{-1}$ ) também foi suficiente para separar os estádios (dados não apresentados), mas a explicação do PC1 foi bastante inferior comparada a 1ª faixa (4.500 a 6.000  $\text{cm}^{-1}$ ). Na 2ª faixa, a variabilidade foi explicada em 63% e 27% para *I. hederifolia* e em 47% e 30%, para *M. aegyptia*, em PC1 e PC2, respectivamente.

Os resultados da PCA para cada uma das espécies reforçam que o comportamento espectral das espécies são distintos e que, embora seja possível estudar modelos mistos (com diferentes grupos de espécies), para PCA os melhores resultados são obtidos quando a análise é aplicada para uma única espécie. Porém, para aplicação prática, o ideal seria conseguir altas acurácias aplicando modelos mistos, visto que, no campo, raramente será encontrada uma única espécie em um único estágio (Souza et al., 2020).

A fim de compreender a segregação das espécies dentro de cada estágio, realizou-se o estudo da PCA para o espectro todo, para as duas faixas somadas e para as faixas individuais. A variabilidade dos dados foi bem explicada para todas as análises, com exceção para o estudo da PCA no espectro todo, em que a variabilidade não foi tão representativa como no estudo das faixas. Para o

espectro todo os três primeiros componentes principais somados, representaram 76%, 59% e 61% para os estádios 1, 2 e 3, respectivamente.



**Figura 5.** Valores de PC 1 e 2 resultantes da análise de componente principal (PCA) na 1ª faixa do espectro (4500 a 6000  $\text{cm}^{-1}$ ) para as espécies de *I.hederifolia* (a) e *M. aegyptia* (b) nos três estádios de crescimento. (triângulo: estágio 1; círculo: estágio 2; quadrado: estágio 3).

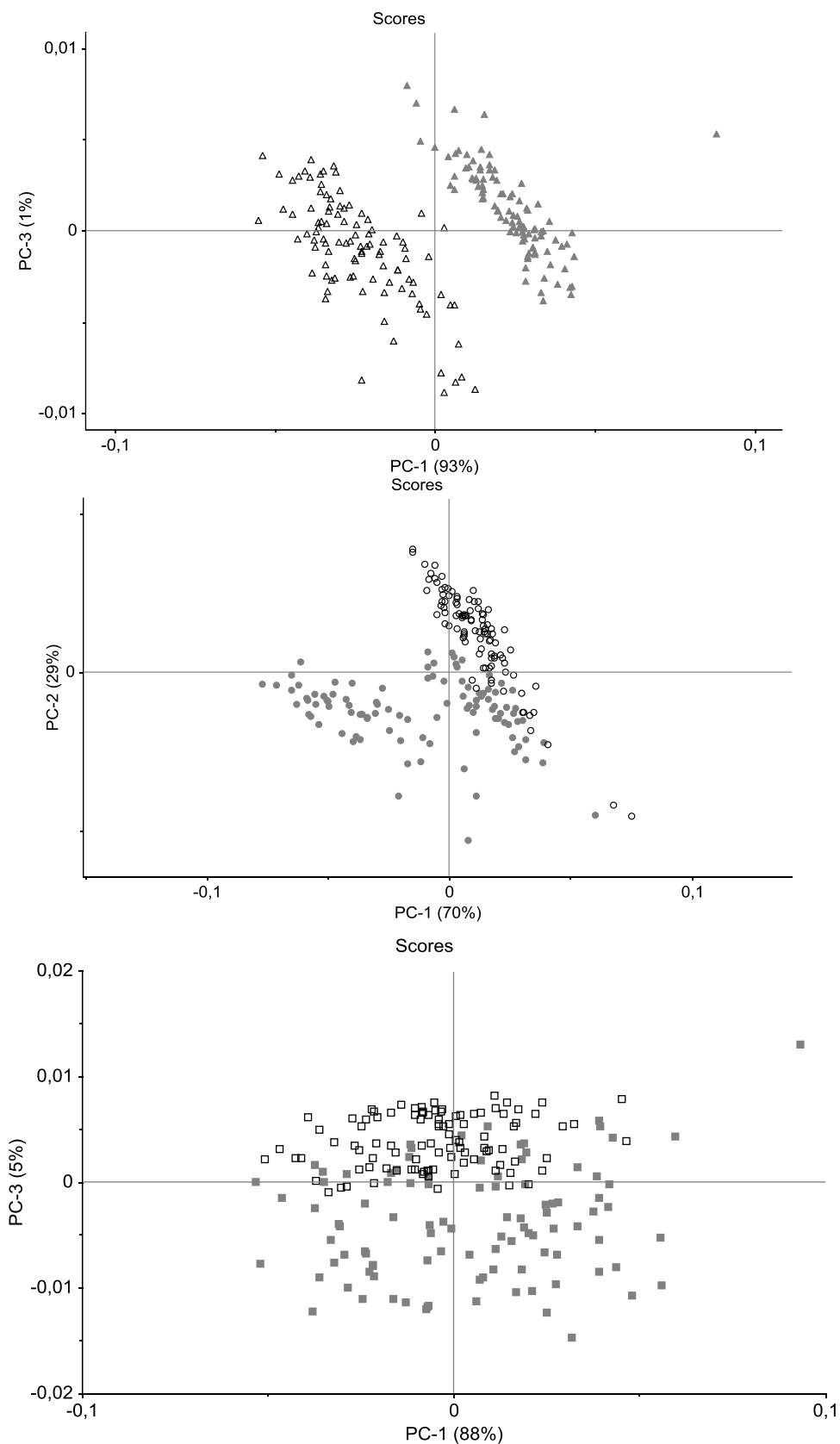
No estudo da 1ª faixa foi possível observar uma melhor separação das espécies, principalmente nos estádios 1 (Figura 6 a) e 2 (Figura 6 b), enquanto no estádio 3 (Figura 6 c) houve maior sobreposição entre as espécies. O menor teor de água observado no estádio 1 (78,93%) em relação aos outros dois estádios (Tabela 2), resultou em uma melhor segregação das espécies. Este resultado ressalta que as diferenças na composição química das espécies se acentuam quando o teor de água é reduzido, contribuindo para um melhor agrupamento das amostras.

O uso da espectroscopia NIR para monitoramento e avaliação da qualidade de produtos agrícolas é bastante empregado em amostras secas ou com baixos teores de umidade, principalmente para melhorar a compreensão da composição química e mitigar a interferência da água no espectro (Cunha Júnior et al., 2015; Izadiyan et al., 2018; Giraudo et al., 2019). Portanto, para PCA aplicada aos dados espectrais de folhas frescas, a segregação das amostras geralmente está associada ao conteúdo de água (León e Downey, 2006; Torres et al., 2019).

Os resultados da PCA confirmam a dificuldade de discriminação das espécies e estádios através de folhas frescas e sugerem que o uso da PCA é mais indicado para amostras com menores teores de umidade. Para tanto, para conseguir resultados mais promissores, o uso de análises supervisionadas, como por exemplo PC-LDA, neste caso, são bastantes comuns (Morais e Lima, 2018; Borraz-Martínez et al., 2019; Souza et al., 2020).

### **3.3. Análise linear discriminante (PC-LDA)**

A análise discriminante linear (PC-LDA) aplicada aos 10 primeiros componentes do espectro todo e das faixas sem pré-processamento foi suficiente para classificar as amostras nas seis classes, correspondentes as duas espécies e aos 3 estádios (Tabela 3). A porcentagem de acerto mínima foi de 90,93% e a máxima de 100%. As faixas com menores erros foram as dos picos somados (1ª faixa + 2ª faixa: 4.500 a 6.000 + 6.500 a 7.750  $\text{cm}^{-1}$ ) e a 1ª faixa (4.500 a 6.000  $\text{cm}^{-1}$ ). Nestas regiões do espectro, o modelo classificou erroneamente duas amostras de *M. aegyptia* do estádio 2 como *I. hederifolia* no estádio 2.



**Figura 6.** Valores dos componentes principais resultantes da análise de componente principal (PCA) para a 1ª faixa do espectro (4500 a 6000  $\text{cm}^{-1}$ ) para as espécies *I. hederifolia* (cinza) e *M. aegyptia* (preto) nos estádios 1(a), 2(b) e 3(c).

**Tabela 3.** Matriz da PC-LDA e porcentagem de acerto para espécies e estádios nas faixas do espectro de 4.000 a 10.000 (espectro todo), 4.500 a 6.000 + 6.500 a 7.750, 4.500 a 6.000 e 6.500 a 7.750  $\text{cm}^{-1}$  sem pré-processamento.

| Espécies e estádios <sup>1</sup>                                                       | IH1 (n=21) | IH2 (n=31) | IH3 (n=35) | MA1 (n=30) | MA2 (n=22) | MA3 (n=41) | Acerto (%) |
|----------------------------------------------------------------------------------------|------------|------------|------------|------------|------------|------------|------------|
| <b>Espectro todo (4.000 a 10.000 <math>\text{cm}^{-1}</math>)</b>                      |            |            |            |            |            |            |            |
| 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     |
| <b>1ª faixa + 2ª faixa (4.500 a 6.000 + 6.500 a 7.750 <math>\text{cm}^{-1}</math>)</b> |            |            |            |            |            |            |            |
| 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     |
| <b>1ª faixa (4.500 a 6.000 <math>\text{cm}^{-1}</math>)</b>                            |            |            |            |            |            |            |            |
| 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     |
| <b>2ª faixa (6.500 a 7.750 <math>\text{cm}^{-1}</math>)</b>                            |            |            |            |            |            |            |            |
| 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      |

<sup>1</sup>Espécies: *I. hederifolia* nos estádios 10 (IH1); 11 (IH2) e 12 (IH3); e *M. aegyptia* nos estádios 11 (MA1); 12 (MA2) e 14 (MA3).

**Tabela 4.** Matriz da PC-LDA e porcentagem de acerto para espécies e estádios nas faixas do espectro de 4.000 a 10.000 (espectro todo), 4.500 a 6.000 + 6.500 a 7.750, 4.500 a 6.000 e 6.500 a 7.750  $\text{cm}^{-1}$  com pré-processamento SNV+1SG.

| Espécies e estádios <sup>1</sup>                                                       | IH1 (n=21) | IH2 (n=31) | IH3 (n=35) | MA1 (n=30) | MA2 (n=22) | MA3 (n=41) | Acerto (%) |
|----------------------------------------------------------------------------------------|------------|------------|------------|------------|------------|------------|------------|
| <b>Espectro todo (4.000 a 10.000 <math>\text{cm}^{-1}</math>)</b>                      |            |            |            |            |            |            |            |
| 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      |
| <b>1ª faixa + 2ª faixa (4.500 a 6.000 + 6.500 a 7.750 <math>\text{cm}^{-1}</math>)</b> |            |            |            |            |            |            |            |
| 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     |
| <b>1ª faixa (4.500 a 6.000 <math>\text{cm}^{-1}</math>)</b>                            |            |            |            |            |            |            |            |
| 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     |
| <b>2ª faixa (6.500 a 7.750 <math>\text{cm}^{-1}</math>)</b>                            |            |            |            |            |            |            |            |
| 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     |

<sup>1</sup>Espécies: *I. hederifolia* nos estádios 10 (IH1); 11 (IH2) e 12 (IH3); e *M. aegyptia* nos estádios 11 (MA1); 12 (MA2) e 14 (MA3).

Ao aplicar PC-LDA nas faixas com o pré-processamento de SNV+1SG (Tabela 4), os acertos variaram entre 40,91% a 96,77%. Para o estudo das faixas, os resultados foram superiores, principalmente para a 2ª faixa (6.500 a 7.750  $\text{cm}^{-1}$ ) que apresentou acertos variando entre 90,91% a 100%.

Embora seja necessária a aplicação de pré-processamentos para as análises não supervisionadas (PCA), quando aplica-se análises supervisionadas (PC-LDA) o uso desses pré-processamentos é dispensável. Sendo assim, para discriminar *I. hederifolia* de *M. aegyptia* e seus estádios utilizando a análise de PC-LDA, recomenda-se o estudo da 1ª faixa (4.500 a 6.000  $\text{cm}^{-1}$ ) do espectro sem pré-processamento.

Os algoritmos baseados em LDA são muito simples em comparação a outros algoritmos encontrados na literatura. Além disso, esses algoritmos reduzem significativamente os dados, acelerando as análises computacionais (Morais e Lima, 2018). Neste contexto, os resultados encontrados neste estudo permitem velocidade e solidez para discriminação das espécies e estádios de corda-de-viola, possibilitando aplicações futuras para o manejo específico dessas espécies.

A PCA agrupou os estádios de crescimento quando aplicada às espécies individualmente e quando aplicada para separar as espécies nos estádios 1 e 2. A análise de PC-LDA classificou corretamente as espécies e estádios com uma porcentagem de acerto de até 100%. Os melhores resultados de discriminação foram observados nos intervalos da 1ª faixa (4.500 a 6.000  $\text{cm}^{-1}$ ) e da 1ª faixa + 2ª faixa (4.500 a 6.000 + 6.500 a 7.750  $\text{cm}^{-1}$ ).

Os resultados apresentados da PC-LDA para espécies e estádios também apontam para a alta sensibilidade que a técnica possui quando combinada às análises multivariadas. Pequenas diferenças nos estádios das espécies foram suficientes para discriminá-las com até 100% de precisão. Desse modo, os próximos estudos para identificação de espécies de plantas daninhas devem se atentar para a construção de modelos ainda mais robustos, com variações ainda maiores, visto que neste estudo as espécies foram conduzidas em condições semi-controladas, sem considerar as adversidades que podem ser encontradas em situações de estresse bióticos e abióticos comuns no campo.

#### 4. CONCLUSÃO

A espectroscopia NIR tem a capacidade de discriminar espécies de *Ipomoea hederifolia* e *Merremia aegyptia* em diferentes estádios de crescimento com uma precisão de até 100%. As melhores taxas de acertos na PC-LDA foram na 1ª faixa (4.500 a 6.000  $\text{cm}^{-1}$ ) e nas faixas somadas (4.500 a 6.000 + 6.500 a 7.750  $\text{cm}^{-1}$ ) para o espectro sem pré-processamento. Já para o espectro com o pré-processamento de SNV+1SG, os melhores resultados foram para a 2ª faixa (6.500 a 7.750  $\text{cm}^{-1}$ ). Este estudo representa um avanço na pesquisa e implantação da tecnologia NIRS para discriminação de espécies de plantas daninhas para o futuro desenvolvimento de equipamentos que auxiliem a adoção e/ou realização de um manejo específico de plantas daninhas, capaz de contribuir com a redução no uso de herbicidas em cultivos.

## 5. REFERÊNCIAS

- 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 Ciencias** 90(1):521–528.
- 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 and Technology** 111:103477.
- Bhullar MS, Walia US, Singh S, Singh M, Jhala AJ (2012) Control of Morningglories (*Ipomoea spp.*) in Sugarcane (*Saccharum spp.*). **Weed Technology**, 26(1):77–82.
- Bleiholder H, Kirfel H, Langelüddeke P, Stauss R (1991) Codificação Unificada dos Estádios Fenológicos de Culturas e Ervas Daninhas. **Pesquisa Agropecuaria Brasileira** 26:1423–1429.
- 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):1–9.
- Cunha Junior L C, 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.
- 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(4):218–233.
- Giraud 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.
- 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 Volumes** 46(10):217–224.
- Hoagland D, Arnon D (1938) The water-culture method for growing plants without soil. **Agricultural Experiment Station Circular 347** (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(2):259–268.

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.

Kennard ARW, Stone LA (1969) Computer Aided Design of Experiments. **American Society for Quality** 11(1):137–148.

Klerkx L, Jakku E, Labarthe P (2019) 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** 90–91(November):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(6):999–1004.

López-Granados F (2011) Weed detection for site-specific weed management: Mapping and real-time approaches. **Weed Research** 51(1):1–11.

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(3):472–481.

Pagnoncelli FB, Trezzi MM, Brum B, Vidal RA, Portes ÁF, Scalcon EL, Machado A (2017) Morning glory species interference on the development and yield of soybeans. **Bragantia** 76(4):470–479.

Pasquini C (2003) Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society** 14(2):198–219.

Piza CST, Nepomuceno MP, Alves PLDCA (2016) Period prior to interference of morning glory in sugarcane. **Científica** 44(4):543.

Ribeiro NM, Torres BA, Ramos SK, dos 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(7):1090–1098.

Rodrigues BN, Almeida FS (2018) **Guia de herbicidas**. 7.ed. Londrina: Produção Independente, v.1.

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.

de Souza MF, Amaral LR, Oliveira SRM, Coutinho MAN, Ferreira Netto C (2020) Spectral differentiation of sugarcane from weeds. **Biosystems Engineering** 190: 41–46.

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.

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(3):276–285.

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(5):277–284.

## CAPÍTULO 4 - Considerações finais

O manejo de plantas daninhas é fundamental para alcance de altas produtividades e corresponde a uma grande parte do custo de produção. A identificação de espécies de plantas daninhas em tempo real para adoção de um manejo específico, representa economia de insumos, menor exposição ambiental e garantia de altas produtividades. Neste contexto, os avanços em tecnologia associados a agricultura de precisão tem possibilitado a realização de estudos sobre a viabilidade do uso de ferramentas auxiliares para um manejo mais específico destas espécies.

Os resultados apresentados neste estudo, bem como as técnicas e análises estudadas, representam um novo olhar para a implementação da espectroscopia no infravermelho próximo (NIR) na área de biologia e manejo de plantas daninhas. A espectroscopia NIR, pode ser utilizada para discriminação de espécies de plantas pertencentes a uma mesma família botânica e até mesmo pertencentes a um único gênero. Além disso, a espectroscopia NIR também foi capaz de discriminar diferentes estádios de crescimento em plantas *in natura*, o que pode auxiliar na escolha de um momento ideal para a realização do controle ou até mesmo, ajuste de doses.

Para discriminação das espécies de *I. hederifolia*, *I. nil* e *M. aegyptia* a acurácia foi de até 97%, indicando uma alta sensibilidade da técnica para diferenciação dessas espécies. Vale ressaltar que os resultados de acurácia dos modelos foram variáveis em função dos pré-processamentos aplicados e também das faixas espectrais analisadas, destacando assim, a importância de estudos exploratórios como este.

Para discriminação das espécies de *I. hederifolia* e *M. aegyptia* considerando os estádios de crescimento, os resultados mostram-se ainda mais sensíveis a técnica, com melhor assertividade. Através da análise não supervisionada de componente principal (PCA) foi possível observar a separação dos grupos de cada espécie e estágio. Ao utilizar a análises supervisionadas, a porcentagem de acerto foi de até 100% nas amostras sem nenhum pré-processamento. Estes resultados demonstram uma alta sensibilidade da técnica para pequenas variações, como os estádios de crescimento. Por isso, é importante investigar as diferenças espectrais em função de

diversas variáveis para uma melhor compreensão e construção de modelos padrões mais robustos.

A espectroscopia NIR é uma técnica promissora para o desenvolvimento de equipamentos capazes de identificar plantas daninhas a campo. Esta técnica quando associada a métodos estatísticos multivariados pode ser usada como uma tecnologia auxiliar no desenvolvimento de sensores e/ou câmeras, baseados na aplicação de tratamentos matemáticos e seleção de faixas espectrais. Trabalhos futuros deverão se concentrar em estudar a influência de outros fatores no comportamento espectral das espécies estudadas e em outras plantas daninhas, assim como, investigar a construção de banco de dados mais robusto, com maiores variações entre os estádios de desenvolvimento, aspectos anatômicos, nutricionais, geográficos e condições de estresse.